LARGE DEVIATIONS PRINCIPLES FOR STOCHASTIC SCALAR CONSERVATION LAWS

MAURO MARIANI

Abstract. Large deviations principles for a family of scalar $1+1$ dimensional conservative stochastic PDEs (viscous conservation laws) are investigated, in the limit of jointly vanishing noise and viscosity. A first large deviations principle is obtained in a space of Young measures. The associated rate functional vanishes on a wide set, the so-called set of measure-valued solutions to the limiting conservation law. A second order large deviations principle is therefore investigated, however, this can be only partially proved. The second order rate functional provides a generalization for non-convex fluxes of the functional introduced by Jensen [12] and Varadhan [21] in a stochastic particles system setting.

1. Introduction

Macroscopic description of physical systems with a large number of degrees of freedom can be often provided by the means of partial differential equations. Rigorous microscopic derivations of such PDEs have been proved in different settings, and we will refer in particular to stochastic interacting particles systems [13, 20], where stochastic microscopic dynamics of particles are considered. One is usually interested in the asymptotic properties of the empirical measures associated with some relevant physical quantities of the system, such as the particles density. Provided that time and space variables are suitably rescaled, it has been proved for several models that, as the number of particles diverges to infinity, the empirical measure associated with the particles density converges to a "macroscopic density" $u \equiv u(t,x)$. Moreover such a density u solves a limiting "hydrodynamical equation", which in the conservative case has usually the following structure

$$\partial_t u + \nabla \cdot \big(f(u) - D(u)\nabla u\big) = 0 \tag{1.1}$$

Here ∇ and $\nabla\cdot$ stands for the space gradient and divergence operators, $D \geq 0$ is a diffusion coefficient, while the flux f takes into account the transport phenomena that may occur in the system. Roughly speaking, D is strictly positive for symmetric (or zero mean) and weakly asymmetric systems, in which case (1.1) is usually obtained in the so-called *diffusive scaling* of the time

and space variables. The case $D \equiv 0$ is instead associated with asymmetric systems, and is usually obtained in the so-called *Euler scaling.*

Once the hydrodynamics of the density is understood, a deeper insight into the system behavior is provided by the investigation of large deviations for the probability law of the empirical measure associated with the density. Establishing large deviations for these models can in fact provide a better understanding of the concepts of entropy and fluctuations in the context of non-equilibrium statistical mechanics. However, while several large deviations results have been obtained for symmetric (or weakly asymmetric) systems under diffusive scaling [13], very little is known for asymmetric systems, with the remarkable exception of the seminal works [15, 12, 21]. According to [13, Chap. 8], large deviations for asymmetric processes are "one of the main open questions in the theory of hydrodynamical limits".

1.1. **Stochastic conservation laws.** In this paper we will focus on a slightly different approach. We consider a continuous "mesoscopic density" $u^\varepsilon \equiv u^\varepsilon(t,x) \in \mathbb{R}$ depending on a small parameter ε (which should be regarded as the inverse of the number of particles). We assume that u^ε satisfies a continuity equation, with a stochastic current taking into account the transport, diffusion and fluctuation phenomena that may occur in the system. More precisely, for $\varepsilon, \gamma > 0$ we consider the stochastic PDE in the unknown u

$$\partial_t u + \nabla \cdot \big(f(u) - \frac{\varepsilon}{2} D(u) \nabla u - \varepsilon^\gamma \sqrt{a^2(u)}\, \alpha^\varepsilon \big) = 0 \tag{1.2}$$

where a^2 is a fluctuation coefficient, and α^ε is a stochastic noise, white in time and with a correlation in space regulated by a convolution kernel $\jmath^\varepsilon$. We assume that $\jmath^\varepsilon$ converges to the identity as $\varepsilon \to 0$, namely that the the range of spatial correlations vanishes at the macroscopic scale. We are then interested in the asymptotic properties (convergence and large deviations) of the solution u^ε to (1.2), as $\varepsilon \to 0$, namely as diffusion and noise vanish *simultaneously.* We remark that, while equations of the form (1.2) may describe quite general physical systems, the limit $\varepsilon \to 0$ is indeed motivated by the heuristic behavior of the density of asymmetric particles systems under Euler scaling. In fact, while one expects the stochastic noise and its spatial correlation to vanish at a macroscopic scale for quite general systems, the limit of jointly vanishing viscosity and noise is somehow specific for the Euler scaling. This specific feature may be one of the (several) reasons making the large deviations of asymmetric systems more challenging.

From the point of view of stochastic PDEs, the limit $\varepsilon \to 0$ also introduces new difficulties. In fact, large deviations for diffusion processes have been widely investigated [10, 7] in the vanishing noise case, and general methods are available to identify the rate functionals associated with large deviations. On the other hand, at our knowledge no results are available -even for finite

dimensional diffusions- if vanishing noise and deterministic drift with nontrivial limiting behavior are considered (here the deterministic drift has a so-called singular limit, see (1.4)). As shown below, in this more general case one needs to investigate a (deterministic) variational problem associated with the stochastic equation. The variational problem associated to (1.2) has been addressed in [3] in a slightly different setting, and we will use most of the results therein obtained.

With respect to the models usually considered in particles systems, (1.2) allows us to get rid of several technicalities related to the discrete nature of particles; we may thus provide a unified treatment of several models (that is, f, D and a are arbitrary). However, as discussed below, the results obtained (namely the speed and rates of large deviations) are in substantial agreement with [12, 21] if the case $f(u) = a^2(u) = u(1-u)$ and $D(u) = 1$ is considered.

1.2. **Outline of the results.** Informally setting $\varepsilon = 0$ in (1.2), we obtain the deterministic PDE

$$\partial_t u + \nabla \cdot f(u) = 0 \tag{1.3}$$

usually referred to as a *conservation law*. As well known [5, Chap. 4], if f is nonlinear, the Cauchy problem associated to (1.3) does not admit global smooth solutions, even if the initial datum is smooth. In general there exist infinitely many weak solutions to (1.3), and an additional *entropic condition* is needed to recover uniqueness and to identify the relevant physical weak solution to (1.3). While (1.3) is invariant under the transformation $(t, x) \mapsto (-t, -x)$, the entropic condition selects a direction of the time, by requiring that entropy is dissipated. A classical result in PDE theory states that the solution to

$$\partial_t u + \nabla \cdot \big(f(u) - \frac{\varepsilon}{2} D(u)\nabla u\big) = 0 \tag{1.4}$$

converges to the entropic solution to (1.3) as $\varepsilon \to 0$, provided the initial data also converge. At the heuristic level, the entropic condition keeps memory of the diffusive term in (1.4) which indeed breaks the symmetry $(t, x) \mapsto (-t, -x)$. We will briefly recall the definition of entropic *Kruzkov* solutions to (1.3) in Section 2, and refer to [5] for an introduction to conservation laws.

There is only a few literature for existence and uniqueness of solutions to fully nonlinear stochastic parabolic equations, see e.g. [16] and [17] dealing with finite-dimensional noise. Under general hypotheses, in the appendix we provide existence and uniqueness (for ε small enough and $\gamma > 1/2$) for the Cauchy problem associated to (1.2), by the means of a piecewise semilinear approximation of such equation. In Section 3.1 we gather some a priori bounds for the solution u^ε to (1.2), and show that, as $\varepsilon \to 0$, u^ε converges in probability to the entropic Kruzkov solution to (1.3) in a strong topology.

We next analyze large deviations principles for the law of u^ε as $\varepsilon \to 0$. In order to avoid technical difficulties associated with the unboundedness of u^ε, and in order to keep our setting as close as possible to the one considered in [12, 21], we assume that the fluctuation coefficient $a^2(u)$ vanishes for $u \not\in (0,1)$. As we will also assume the initial datum to take values in $[0,1]$, this condition guarantees that u^ε takes values in $[0,1]$, see Theorem A.1. We only consider the $(1+1)$ dimensional case, with the (t,x) variables running in $[0,T]\times\mathbb{T}$, where $T>0$ and $\mathbb{T}$ is the one dimensional torus. While these restrictions are merely technical, we remark that only the case of scalar u is considered, as the vectorial case (systems of conservation laws) is certainly far more difficult.

In Section 3.2 we establish a large deviations principle with speed $\varepsilon^{-2\gamma}$, roughly equivalent to the classical Freidlin-Wentzell speed for finite dimensional diffusions [10]. The bottom line is that, when events with probability of order $e^{\varepsilon^{-2\gamma}}$ are considered, the noise term in (1.2) can bitterly deviate from its "typical behavior" thus completely overcoming the regularizing effect of the vanishing parabolic term. Any entropy-dissipation phenomena is lost at this speed, and the noise may drive severe oscillations of the density u^ε as $\varepsilon \to 0$. The large deviations are then naturally investigated in a Young measures setting. We prove that on a Young measure $\mu \equiv \mu_{t,x}(d\lambda)$ (satisfying a suitable initial condition) the large deviations rate functional is given by (see Section 2.4 for a more precise definition of $\mathcal{I}$)

$$\mathcal{I}(\mu) := \frac{1}{2}\int_0^T dt\, \Big\| \partial_t \mu(\imath) + \nabla\cdot\mu(f)\Big\|^2_{H^{-1}(\mu(a^2),dx)}$$

Here $\imath:\mathbb{R}\to\mathbb{R}$ is the identity map, for F a continuous function, $\mu(F)(t,x)$ stands for $\int \mu_{t,x}(d\lambda)F(\lambda)$, and with a little abuse of notation, we denoted by $\|\varphi\|_{H^{-1}(\mu_{t,\cdot}(a^2),dx)}$ the dual norm to $\big[\int dx\,\mu_{t,x}(a^2)\,\varphi_x^2\big]^{1/2}$.

Note that $\mathcal{I}(\mu)=0$ iff μ is a measure-valued solution to (1.3) (see Section 2.4). The Cauchy problem (1.3) admits in general infinitely many measure-valued solutions, but we stated above that u^ε converges in probability to the (unique) entropic solution to (1.3). One thus expects that nontrivial large deviations principle may hold with a speed slower than $\varepsilon^{-2\gamma}$. In Section 3.3, we investigate large deviations principle with speed $\varepsilon^{-2\gamma+1}$. At this scale, deviations of the noise term in (1.3) are of the same order of the parabolic term. The law of u^ε is then exponentially tight (with speed $\varepsilon^{-2\gamma+1}$) in a suitable space of functions. To informally define the candidate rate functional for the large deviations with this speed, we briefly introduce some preliminary notions, which will be precisely explained in Section 2.5.

We say that a weak solution u to (1.3) is an *entropy-measure* solution iff there exists a measurable map ϱ_u from $[0,1]$ to the set of Radon measures on

$(0,T) \times \mathbb{T}$, such that for each $\eta \in C^2([0,1])$ and $\varphi \in C_c^\infty\big((0,T)\times\mathbb{T}\big)$

$$-\int dt\, dx\, \big[\eta(u)\varphi_t + q(u)\nabla\varphi\big] = \int dv\, \varrho_u(v;dt,dx)\eta''(v)\varphi(t,x)$$

where $q(v) := \int^v dw\, \eta'(w)f'(w)$, see Proposition 2.6 for a characterization of entropy-measure solutions to (1.3). The candidate rate functional for the second order large deviations is the functional H defined as follows. If u is not an entropy-measure solution to (1.3) then $H(u) = +\infty$. Otherwise $H(u) = \int dv\, \varrho_u^+(v;dt,dx)D(v)\, a^{-2}(v)$, where ϱ_u^+ denotes the positive part of ϱ_u. Note that H depends on the diffusion coefficient D and the fluctuation coefficient a^2 only through their ratio, thus fitting in the *Einsten paradigm* for macroscopic diffusive systems. We also remark that, while the functional $\mathcal{I}$ is convex, H is not (for instance, convex combinations of entropy-measure solutions to (1.3), in general are not weak solutions).

While we prove a large deviations upper bound with speed $\varepsilon^{-2\gamma+1}$ and rate H, we obtain the lower bound only on a suitable set $\mathcal{S}$ of weak solutions to (1.3), see Definition 2.7. To complete the proof of this *second order* large deviations, an additional density argument is needed. This seems to be a challenging problem, and as noted by Varadhan in [21] "…one does not see at the moment how to produce a 'general' non-entropic solution, partly because one does not know what it is."

It is easy to see that, on the set of weak solutions to (1.3) with bounded variations and on the set $\mathcal{S}$, the rate functional H^{JV} introduced in [12, 21] coincides with the rate functional H evaluated for $f(v) = v(1-v)$, $D \equiv 1$ and $a^2(v) = v(1-v)$, which are the expected transport, diffusion and fluctuation coefficients for the totally asymmetric simple exclusion process there investigated. In particular, H comes as a natural generalization of the functional introduced in [12, 21], whenever the flux f is neither convex nor concave. Unfortunately, since chain rule formulas are not available out of the BV setting, one cannot check that $H = H^{JV}$ on the whole set of entropy-measure solutions to (1.3). Note however that the inequality $H \geq H^{JV}$ holds. Furthermore, under smoothness and genuine nonlinearity assumption on f, $H(u) = 0$ iff u is the unique entropic solution to (1.3), so that higher order large deviations principles are trivial.

1.3. **Outline of the proof.** The convergence in probability of u^ε to the entropic solution of (1.3) is obtained by a sharp stability analysis of the stochastic perturbation (1.2) of (1.4).

The large deviations upper bound with speed $\varepsilon^{-2\gamma}$ is provided by lifting the standard Varadhan's minimax method to the Young measures setting, while exponential tightness in this space is easily proved. The corresponding lower bound is first proved for Young measures that are Dirac masses at almost every

point $(t,x) \in [0,T]\times\mathbb{T}$, and then extended to the whole set of Young measures by adapting the relaxation argument in [3].

The large deviations with speed $\varepsilon^{-2\gamma+1}$ are much different than the usual small noise asymptotic limit for Itô processes. Note indeed that, as $\varepsilon \to 0$, the parabolic term in (1.3) has a nontrivial behavior. In such a case there is no general method to study large deviations, even in a finite dimensional setting. We provide a link of the large deviations problem with a Γ-convergence result obtained in [3]. Indeed we use the equicoercivity of a suitable family of functionals to show exponential tightness, and we use the so-called Γ-limsup result to build up the optimal exponential martingales for the lower bound. In particular, since the Γ-limsup inequality in [3] is not fully established, we only have partial results for the lower bound. The upper bound is established by a nonlinear version of the Varadhan's minimax method.

2. Main results

2.1. **Notation.** In this paper, $T > 0$ is a positive real number and we let $\big(\Omega, \mathfrak{F}, \{\mathfrak{F}_t\}_{0\le t\le T}, P\big)$ be a *standard* filtered probability space. For B a real Banach space and $M : [0,T]\times\Omega \to B$ a given adapted process, we write equivalently $M(t) \equiv M(t,\omega)$. For each $\phi \in B^*$ we denote by $\langle M,\phi\rangle \equiv \langle M,\phi\rangle(t,\omega)$ the real–valued process obtained by the dual action of M on B. Given two real–valued P-square integrable martingales M, N, we denote by $\big[M,N\big] \equiv \big[M,N\big](t,\omega)$ the cross quadratic variation process of M and N. In the following *martingale* will always stand for *continuous martingale*. For a Polish space X, we also let $\mathcal{P}(X)$ denote the set of Borel probability measures on X. For ν a measure on some measurable space and $F \in L_1(d\nu)$, we denote by $\nu(F)$ the integral of F with respect to ν. However, for a probability P we used the notation $\mathbb{E}^P$ to denote the expected value.

We denote by $\mathbb{T}$ the one-dimensional torus, by $\langle\cdot,\cdot\rangle$ the inner product in $L_2(\mathbb{T})$, and by $\langle\langle\cdot,\cdot\rangle\rangle$ the inner product in $L_2([0,T]\times\mathbb{T})$. For E a closed set in $[0,T]\times\mathbb{T}$, $C^k(E)$ denotes the collection of k-times differentiable functions on E, with continuous derivatives up to the boundary. We also let $H^1(\mathbb{T})$ be the Hilbert space of square integrable functions on $\mathbb{T}$ with square integrable derivative, and let $H_{-1}(\mathbb{T})$ be its dual space. Throughout this paper ∂_t denotes derivative with respect to the time variable t, ∇ and $\nabla\cdot$ derivatives with respect to the space variable x (while we consider a one dimensional space setting, we consider gradient and divergence as distinct operators). For a function ϑ explicitly depending on the x variable, ∂_x denotes the partial derivative with respect to x. Namely, given a function $u : \mathbb{T} \to [0,1]$ and $\vartheta : [0,1]\times\mathbb{T} \to \mathbb{R}$, we understand $\nabla[\vartheta(u(x),x)] = (\partial_u\vartheta)(u(x),x)\nabla u(x) + (\partial_x\vartheta)(u(x),x)$. In the following we will usually omit the dependence on the ω variable, as well as on the t and/or x variables when no misunderstanding is possible.

2.2. **Stochastic conservation laws.** We refer to [7] for a general theory of stochastic differential equations in infinite dimensions. Let W be an $L_2(\mathbb{T})$–valued cylindrical Brownian motion on $\big(\Omega, \mathfrak{F}, \{\mathfrak{F}_t\}_{0\le t\le T}, P\big)$. Namely, W is a Gaussian, $L_2(\mathbb{T})$–valued P-martingale with quadratic variation:

$$\big[\langle W,\phi\rangle, \langle W,\psi\rangle\big](t,\omega) = \langle \phi,\psi\rangle\, t \tag{2.1}$$

for each $\phi,\psi \in L_2(\mathbb{T})$.

For $\varepsilon > 0$, we consider the following stochastic Cauchy problem in the unknown u:

$$\begin{aligned} du &= \Big[-\nabla\cdot f(u) + \frac{\varepsilon}{2}\nabla\cdot\big(D(u)\nabla u\big)\Big]\, dt + \varepsilon^{\gamma}\,\nabla\cdot\big[a(u)(\jmath^{\varepsilon} * dW)\big] \\ u(0,x) &= u_0^{\varepsilon}(x) \end{aligned} \tag{2.2}$$

Here $\gamma > 0$ is a real parameter, and $\nabla\cdot\big[a(u)(\jmath^{\varepsilon} * dW)\big]$ stands for the martingale differential acting on $\psi \in H^1(\mathbb{T})$ as

$$\big\langle \nabla\cdot\big[a(u)(\jmath^{\varepsilon} * dW)\big], \psi\big\rangle = -\langle dW, \jmath^{\varepsilon} * [a(u)\nabla\psi]\rangle$$

The following hypotheses will be always assumed below, but in the appendix.

H1) $f : [0,1] \to \mathbb{R}$ is a Lipschitz function.
H2) $D : [0,1] \to \mathbb{R}$ is a uniformly positive Lipschitz function.
H3) $a \in C^2([0,1])$ is such that $a(0) = a(1) = 0$, and $a(v) \neq 0$ for $v \in (0,1)$.
H4) $\{\jmath^{\varepsilon}\}_{\varepsilon>0} \subset H^1(\mathbb{T})$ is a sequence of positive mollifiers with $\int dx\, \jmath^{\varepsilon}(x) = 1$, weakly converging to the Dirac mass centered at 0.
H5) For $\varepsilon > 0$, $u_0^{\varepsilon} : \Omega\times\mathbb{T} \to [0,1]$ is a measurable map with respect to the product $\mathfrak{F}_0 \times$ Borel σ-algebra. Moreover there exists a Borel measurable function $u_0 : \mathbb{T} \to [0,1]$ such that, for each $\delta > 0$

$$\lim_{\varepsilon} P\big(\|u_0^{\varepsilon} - u_0\|_{L_1(\mathbb{T})} > \delta\big) = 0$$

The next proposition is an immediate consequence of Proposition A.7 in the appendix, where we also recall the precise definitions of strong and martingale solutions to (2.2) and we briefly discuss why the condition on γ and $\jmath^{\varepsilon}$ (see Proposition 2.1 below) are needed.

Proposition 2.1. *Assume* $\lim_{\varepsilon} \varepsilon^{2\gamma-1}\|\jmath^{\varepsilon}\|^2_{L_2(\mathbb{T})} = 0$. *Then there is an* $\varepsilon_0 > 0$ *depending only on* D *and* a, *such that, for each* $\varepsilon < \varepsilon_0$, *there exists a unique adapted process* $u^{\varepsilon} : \Omega \to C\big([0,T]; H^{-1}(\mathbb{T})\big) \cap L_2\big([0,T]; H^1(\mathbb{T})\big)$ *solving* (2.2) *in the strong stochastic sense. Moreover* u^{ε} *admits a version in* $C\big([0,T]; L_1(\mathbb{T})\big)$, *and for every* $t \in [0,T]$ $u^{\varepsilon}(\omega; t, x) \in [0,1]$ *for* $dP\, dx$ *a.e.* (ω, x).

Note that the total mass of u^{ε} is conserved a.s. by the stochastic flow (2.2), namely for each $t \in [0,T]$ we have $\int dx\, u^{\varepsilon}(t,x) = \int dx\, u_0^{\varepsilon}(x)$ P a.s.. We are interested in the asymptotic limit of the probability law of the solution u^{ε} to (2.2) as $\varepsilon \to 0$.

2.3. **Deterministic conservation laws.** Let U denote the compact Polish space of measurable functions $u : \mathbb{T} \to [0,1]$, equipped with the metric it inherits as a (closed) subset of $H^{-1}(\mathbb{T})$, namely

$$d_U(u,v) := \sup \Big\{ \langle u - v, \varphi \rangle,\, \varphi \in H^1(\mathbb{T}) \,:\, \|\varphi\|^2_{L_2(\mathbb{T})} + \|\nabla \varphi\|^2_{L_2(\mathbb{T})} \le 1 \Big\}$$

Fix $T > 0$ and consider the formal limiting equation for (2.2)

$$\begin{aligned} &\partial_t u + \nabla \cdot f(u) = 0 \\ &u(0,x) = u_0(x) \end{aligned} \tag{2.3}$$

In general there exist no smooth solutions to (2.3). A function $u \in C\big([0,T]; U\big)$ is a *weak solution* to (2.3) iff for each $\varphi \in C^\infty([0,T] \times \mathbb{T})$ it satisfies

$$\langle u(T), \varphi(T) \rangle - \langle u_0, \varphi(0) \rangle - \langle\langle u, \partial_t \varphi \rangle\rangle - \langle\langle f(u), \nabla \varphi \rangle\rangle = 0$$

As well known [5, Chap. 6], existence and uniqueness of a weak *Kruzkov* solution to (2.3) is guaranteed under an additional entropic condition, which is recalled in Section 2.5 below. Then u^ε converges in probability to such a solution both in the strong $L_p([0,T] \times \mathbb{T})$ and $C\big([0,T]; U\big)$ topologies.

Proposition 2.2. *Assume* $\lim_\varepsilon \varepsilon^{2(\gamma-1)} \big[\|\jmath^\varepsilon\|^2_{L_2(\mathbb{T})} + \varepsilon \|\nabla \jmath^\varepsilon\|^2_{L_2(\mathbb{T})} \big] = 0$. *Let* $\bar{u}$ *be the unique Kruzkov solution to* (2.3). *Then for each* $p < +\infty$ *and* $\delta > 0$

$$\lim_\varepsilon P\Big(\|u^\varepsilon - \bar{u}\|^p_{L_p([0,T] \times \mathbb{T})} + \sup_{t \in [0,T]} d_U(u^\varepsilon(t), \bar{u}(t)) > \delta \Big) = 0$$

Proposition 2.2 establishes a convergence result for the probability law of the process u^ε solution to (2.2), as $\varepsilon \to 0$. We are then interested in large deviations principles for this probability law. We recall the definition of the large deviations bounds [8].

Definition 2.3. *Let* $\mathcal{X}$ *be a Polish space and* $\{\mathbb{P}^\varepsilon\} \subset \mathcal{P}(\mathcal{X})$ *a family of Borel probability measures on* $\mathcal{X}$. *For* $\{\alpha_\varepsilon\}$ *a sequence of positive reals such that* $\lim_\varepsilon \alpha_\varepsilon = 0$ *and* $I : \mathcal{X} \to [0, +\infty]$ *a lower semicontinuous functional, we say that* $\{\mathbb{P}^\varepsilon\}$ *satisfies*

- *A* large deviations upper bound *with speed* α_ε^{-1} *and rate* I, *iff for each closed set* $\mathcal{C} \subset \mathcal{X}$

$$\overline{\lim_\varepsilon}\, \alpha_\varepsilon \log \mathbb{P}^\varepsilon(\mathcal{C}) \le - \inf_{u \in \mathcal{C}} I(u) \tag{2.4}$$

- *A* large deviations lower bound *with speed* α_ε^{-1} *and rate* I, *iff for each open set* $\mathcal{O} \subset \mathcal{X}$

$$\overline{\lim_\varepsilon}\, \alpha_\varepsilon \log \mathbb{P}^\varepsilon(\mathcal{O}) \ge - \inf_{u \in \mathcal{O}} I(u) \tag{2.5}$$

$\{\mathbb{P}^\varepsilon\}$ *is said to satisfy a* large deviations principle *if both the upper and lower bounds hold with same rate and speed.*

In the next sections, we introduce some preliminary notions and state a first large deviations principle with speed $\varepsilon^{-2\gamma}$. We next introduce some additional preliminaries and state a second large deviations partial result, associated with the speed $\varepsilon^{-2\gamma+1}$.

2.4. **First order large deviations.** We first introduce a suitable space $\mathcal{M}$ of Young measures and recall the notion of measure-valued solution to (2.3). Consider the set $\mathcal{N}$ of measurable maps μ from $[0,T]\times\mathbb{T}$ to the set $\mathcal{P}([0,1])$ of Borel probability measures on $[0,1]$. The set $\mathcal{N}$ can be identified with the set of positive finite Borel measures μ on $[0,T]\times\mathbb{T}\times[0,1]$ such that $\mu(dt,\,dx,\,[0,1]) = dt\,dx$, by the bijection $\mu(dt,\,dx,\,d\lambda) = dt\,dx\,\mu_{t,x}(d\lambda)$. For $\imath:[0,1]\to[0,1]$ the identity map, we set

$$\mathcal{M} := \big\{\mu\in\mathcal{N}\,:\,\text{the map }[0,T]\ni t\mapsto\mu_{t,\cdot}(\imath)\text{ is in }C\big([0,T];U\big)\big\}$$

in which, for a bounded measurable function $F:[0,1]\to\mathbb{R}$, the notation $\mu_{t,x}(F)$ stands for $\int_{[0,1]}\mu_{t,x}(d\lambda)F(\lambda)$. We endow $\mathcal{M}$ with the metric

$$d_{\mathcal{M}}(\mu,\nu) := d_{*\mathrm{w}}(\mu,\nu) + \sup_{t\in[0,T]} d_U\big(\mu_{t,\cdot}(\imath),\nu_{t,\cdot}(\imath)\big)$$

where $d_{*\mathrm{w}}$ is a distance generating the relative topology on $\mathcal{N}$ regarded as a subset of the finite Borel measures on $[0,T]\times\mathbb{T}\times[0,1]$ equipped with the $*$-weak topology. $(\mathcal{M},d_{\mathcal{M}})$ is a Polish space.

An element $\mu\in\mathcal{M}$ is a *measure-valued solution* to (2.3) iff for each $\varphi\in C^\infty\big([0,T]\times\mathbb{T}\big)$ it satisfies

$$\langle\mu_{T,\cdot}(\imath),\varphi(T)\rangle - \langle u_0,\varphi(0)\rangle - \langle\langle\mu(\imath),\partial_t\varphi\rangle\rangle - \langle\langle\mu(f),\nabla\varphi\rangle\rangle = 0$$

If $u\in C\big([0,T];U\big)$ is a weak solution to (2.3), then the map $(t,x)\mapsto\delta_{u(t,x)}(d\lambda)\in\mathcal{P}([0,1])$ is a measure-valued solution. However, in general there exist measure-valued solutions which do not have this form, namely they are not a Dirac mass at a.e. (t,x) (e.g. finite convex combinations of Dirac masses centered on weak solutions are measure-valued solutions).

Consider the process $\mu^\varepsilon:\Omega\to\mathcal{M}$ defined by $\mu^\varepsilon_{t,x} := \delta_{u^\varepsilon(t,x)}$. We let $\mathbf{P}^\varepsilon := P\circ(\mu^\varepsilon)^{-1}\in\mathcal{P}(\mathcal{M})$ be the law of μ^ε on $\mathcal{M}$. In Section 3.2 we prove

Theorem 2.4. *Assume* $\lim_\varepsilon \varepsilon^{2(\gamma-1)}\big[\|\jmath^\varepsilon\|^2_{L_2} + \varepsilon\|\nabla\jmath^\varepsilon\|^2_{L_2}\big] = 0$.

(i) *Then the sequence $\{\mathbf{P}^\varepsilon\}\subset\mathcal{P}(\mathcal{M})$ satisfies a large deviations upper bound on $\mathcal{M}$ with speed $\varepsilon^{-2\gamma}$ and rate functional $\mathcal{I}:\mathcal{M}\to[0,+\infty]$ defined as*

$$\mathcal{I}(\mu) := \sup_{\varphi\in C^\infty([0,T]\times\mathbb{T})}\Big\{\langle\mu_{T,\cdot}(\imath),\varphi(T)\rangle - \langle u_0,\varphi(0)\rangle - \langle\langle\mu(\imath),\partial_t\varphi\rangle\rangle - \langle\langle\mu(f),\nabla\varphi\rangle\rangle - \frac{1}{2}\,\langle\langle\mu(a^2)\nabla\varphi,\nabla\varphi\rangle\rangle\Big\} \quad (2.6)$$

(ii) *Assume furthermore that* $\zeta \le u_0 \le 1-\zeta$ *for some* $\zeta > 0$. *Then* $\{\mathbf{P}^\varepsilon\} \subset \mathcal{P}(\mathcal{M})$ *satisfies a large deviations lower bound on* $\mathcal{M}$ *with speed* $\varepsilon^{-2\gamma}$ *and rate functional* $\mathcal{I}$.

We denote by $\mathbb{P}^\varepsilon := P \circ (u^\varepsilon)^{-1} \in \mathcal{P}\big(C\big([0,T];U\big)\big)$ the law of u^ε on the Polish space $(C\big([0,T];U\big)$. By contraction principle [8, Theorem 4.2.1] we get

Corollary 2.5. *Under the same hypotheses of Theorem 2.4, the sequence* $\{\mathbb{P}^\varepsilon\} \subset \mathcal{P}\big(C\big([0,T];U\big)\big)$ *satisfies a large deviations principle on* $C\big([0,T];U\big)$ *with speed* $\varepsilon^{-2\gamma}$ *and rate functional* $I : C\big([0,T];U\big) \to [0,+\infty]$ *defined as*

$$I(u) := \inf\Big\{ \int dt\, dx\, R_{f,a^2}\big(u(t,x), \Phi(t,x)\big),$$
$$\Phi \in L_2([0,T]\times\mathbb{T}) \,:\, \nabla\Phi = -\partial_t u \textit{ weakly}\Big\}$$

where $R_{f,a^2} : [0,1]\times\mathbb{R} \to [0,+\infty]$ *is defined by*

$$R_{f,a^2}(w,c) := \inf\{\big(\nu(f)-c\big)^2/\nu(a^2),\, \nu \in \mathcal{P}([0,1]) \,:\, \nu(\imath) = w\}$$

in which we understand $(c-c)^2/0 = 0$.

Note that, if $\mathcal{I}(\mu) < +\infty$, then $\mu_{0,x}(\imath) = u_0(x)$ and analogously $I(u) < +\infty$ implies $u(0,x) = u_0(x)$. On the other hand, $\mathcal{I}(\mu) = 0$ iff μ is a measure-valued solution to (2.3). $\mathcal{I}(\mu)$ quantifies indeed how μ deviates from being a measure-valued solution to (2.3) in a suitable Hilbert norm, see the proof of Theorem 2.4 item (i) in Section 3.2. On the other hand, if f is nonlinear, in general we have $I(u) < \mathcal{I}(\delta_u)$, so that I vanishes on a set wider than the set of weak solutions to (2.3).

In general there exist infinitely many measure-valued solutions to (2.3), but Proposition 2.2 implies that $\{\mathbf{P}^\varepsilon\}$ converges in probability in $\mathcal{M}$ to the unique Kruzkov solution $\bar{u}$ to (2.3) (more precisely, to the Young measure $\bar{\mu}$ defined by $\bar{\mu}_{t,x} = \delta_{\bar{u}(t,x)}$). We thus expect that additional nontrivial large deviations principles may hold with a speed slower than $\varepsilon^{-2\gamma}$.

2.5. **Entropy-measure solutions to conservation laws.** Let $\mathcal{X}$ be the same set $C([0,T];U)$ endowed with the metric

$$d_{\mathcal{X}}(u,v) := \|u-v\|_{L_1([0,T]\times\mathbb{T})} + \sup_{t\in[0,T]} d_U(u(t),v(t))$$

Convergence in $\mathcal{X}$ is of course strictly stronger than convergence in $C\big([0,T];U\big)$, since convergence in $L_p([0,T]\times\mathbb{T})$ for $p \in [1,+\infty)$ is also required. Note that $\mathcal{X}$ can be identified with the subset of $\mathcal{M}$

$$\big\{\mu \in \mathcal{M} \,:\, \mu = \delta_u, \text{ for some } u \in C\big([0,T];U\big)\big\}$$

and $d_{\mathcal{X}}$ is indeed a distance generating the relative topology induced by $d_{\mathcal{M}}$ on $\mathcal{X}$. In particular, once exponential tightness is established on $\mathcal{X}$, it is

immediate to *lift* large deviations principles for the law of u^ε on $\mathcal{X}$, to the corresponding law of δ_{u^ε} on $\mathcal{M}$.

A function $\eta \in C^2([0,1])$ is called an *entropy* and its *conjugated entropy flux* $q \in C([0,1])$ is defined up to a constant by $q(u) := \int^u dv\, \eta'(v) f'(v)$. For u a weak solution to (2.3), for (η, q) an entropy–entropy flux pair, the η*-entropy production* is the distribution $\wp_{\eta,u}$ acting on $C_c^\infty\big([0,T) \times \mathbb{T}\big)$ as

$$\wp_{\eta,u}(\varphi) := -\langle \eta(u_0), \varphi(0)\rangle - \langle\langle \eta(u), \partial_t \varphi\rangle\rangle - \langle\langle q(u), \nabla\varphi\rangle\rangle \tag{2.7}$$

Let $C_c^{2,\infty}\big([0,1] \times [0,T) \times \mathbb{T}\big)$ be the set of compactly supported maps $\vartheta : [0,1] \times [0,T) \times \mathbb{T} \ni (v,t,x) \to \vartheta(v,t,x) \in \mathbb{R}$, that are twice differentiable in the v variable, with derivatives continuous up to the boundary of $[0,1] \times [0,T) \times \mathbb{T}$, and that are infinitely differentiable in the (t,x) variables. For $\vartheta \in C_c^{2,\infty}\big([0,1] \times [0,T) \times \mathbb{T}\big)$ we denote by ϑ' and ϑ'' its partial derivatives with respect to the v variable. We say that a function $\vartheta \in C_c^{2,\infty}\big([0,1] \times [0,T) \times \mathbb{T}\big)$ is an *entropy sampler*, and its *conjugated entropy flux sampler* $Q : [0,1] \times [0,T) \times \mathbb{T}$ is defined up to an additive function of (t,x) by $Q(u,t,x) := \int^u dv\, \vartheta'(v,t,x) f'(v)$. Finally, given a weak solution u to (2.3), the ϑ*-sampled entropy production* $P_{\vartheta,u}$ is the real number

$$\begin{aligned} P_{\vartheta,u} \;\; := \;\; & -\int dx\, \vartheta(u_0(x), 0, x) \\ & -\int dt\, dx\, \Big[\big(\partial_t \vartheta\big)\big(u(t,x), t, x\big) + \big(\partial_x Q\big)\big(u(t,x), t, x\big)\Big] \end{aligned} \tag{2.8}$$

If $\vartheta(v,t,x) = \eta(v)\varphi(t,x)$ for some entropy η and some $\varphi \in C_c^\infty\big([0,T) \times \mathbb{T}\big)$, then $P_{\vartheta,u} = \wp_{\eta,u}(\varphi)$.

We next introduce a suitable class of solutions to (2.3) for later use. We denote by $M\big([0,T) \times \mathbb{T}\big)$ the set of Radon measures on $[0,T) \times \mathbb{T}$ that we consider equipped with the vague topology. In the following, for $\wp \in M\big([0,T) \times \mathbb{T}\big)$ we denote by $\wp^\pm$ the positive and negative part of $\wp$. For u a weak solution to (2.3) and η an entropy, recalling (2.7) we set

$$\|\wp_{\eta,u}\|_{\mathrm{TV}} := \sup\big\{\wp_{\eta,u}(\varphi),\, \varphi \in C_c^\infty\big([0,T) \times \mathbb{T}\big),\, |\varphi| \le 1\big\}$$

$$\|\wp_{\eta,u}^+\|_{\mathrm{TV}} := \sup\big\{\wp_{\eta,u}(\varphi),\, \varphi \in C_c^\infty\big([0,T) \times \mathbb{T}\big),\, 0 \le \varphi \le 1\big\}$$

The following result follows by adapting [3, Prop. 2.3] and [6, Prop. 3.1] to the setting of this paper.

Proposition 2.6. *Let $u \in \mathcal{X}$ be a weak solution to* (2.3)*. The following statements are equivalent:*

(i) *For each entropy η, the η-entropy production $\wp_{\eta,u}$ can be extended to a Radon measure on $[0,T) \times \mathbb{T}$, namely $\|\wp_{\eta,u}\|_{\mathrm{TV}} < +\infty$ for each entropy η.*

(ii) *There exists a bounded measurable map* $\varrho_u : [0,1] \ni v \to \varrho_u(v; dt, dx) \in M\big([0,T) \times \mathbb{T}\big)$ *such that for any entropy sampler* ϑ

$$P_{\vartheta,u} = \int dv\, \varrho_u(v; dt, dx)\, \vartheta''(v, t, x)$$

A weak solution $u \in \mathcal{X}$ that satisfies the equivalent conditions in Proposition 2.6 is called an *entropy-measure solution* to (2.3). We denote by $\mathcal{E} \subset \mathcal{X}$ the set of entropy-measure solutions to (2.3).

A weak solution $u \in \mathcal{X}$ to (2.3) is called an *entropic solution* iff for each convex entropy η the inequality $\wp_{\eta,u} \le 0$ holds in distribution sense, namely $\|\wp^+_{\eta,u}\|_{\mathrm{TV}} = 0$. Entropic solutions are entropy-measure solutions such that $\varrho_u(v; dt, dx)$ is a negative Radon measure for each $v \in [0,1]$. It is well known, see e.g. [5, Theorem 6.2.1], that for each $u_0 \in U$ there exists a unique entropic weak solution $\bar{u} \in \mathcal{X} \cap C\big([0,T]; L_1(\mathbb{T})\big)$ to (2.3). Such a solution is called the *Kruzkov solution* with initial datum u_0.

Up to minor adaptations, the following class of solutions have been also introduced in [3], where some examples of such solutions are also given.

Definition 2.7. *An entropy-measure solution* $u \in \mathcal{E}$ *is* entropy-splittable *iff there exist two closed sets* $E^+, E^- \subset [0,T] \times \mathbb{T}$ *such that*

(i) *For a.e.* $v \in [0,1]$*, the support of* $\varrho_u^+(v; dt, dx)$ *is contained in* E^+*, and the support of* $\varrho_u^-(v; dt, dx)$ *is contained in* E^-*.*
(ii) *The set* $\big\{t \in [0,T] \,:\, \big(\{t\} \times \mathbb{T}\big) \cap E^+ \cap E^- \neq \emptyset\big\}$ *is nowhere dense in* $[0,T]$*.*
(iii) *There exists* $\delta > 0$ *such that* $\delta \le u \le 1 - \delta$*.*

The set of entropy-splittable solutions to (2.3) *is denoted by* $\mathcal{S}$*.*

Note that $\mathcal{S} \subset \mathcal{E} \subset \mathcal{X}$, and if u_0 is bounded away from 0, 1, then $\mathcal{S}$ is nonempty (for instance the Kruzkov solution to (1.3) is in $\mathcal{S}$). Indeed in general $\mathcal{S} \not\subset BV\big([0,T] \times \mathbb{T}\big)$.

2.6. **Second order large deviations.** With a little abuse of notation, we still denote with $\mathbb{P}^\varepsilon := P \circ (u^\varepsilon)^{-1} \in \mathcal{P}(\mathcal{X})$ the law of u^ε on the Polish space $(\mathcal{X}, d_{\mathcal{X}})$. Since $\int dx\, \jmath^\varepsilon(x) = 1$ (see hypothesis **H4)**), we have that $\jmath^\varepsilon - 1$ is the derivative of some smooth function J on $\mathbb{T}$, defined up to an additive constant. We define $\|\jmath^\varepsilon - \mathbb{1}\|_{W^{-1,1}(\mathbb{T})}$ as the infimum of $\|J\|_{L_1(\mathbb{T})}$ as J runs on the set of functions J such that $\nabla \cdot J = \jmath^\varepsilon - 1$. We have the following

Theorem 2.8. *Assume that there is no interval in* $[0,1]$ *where* f *is affine, and that* $\lim_\varepsilon \varepsilon^{2(\gamma-1)}\big[\|\jmath^\varepsilon\|^2_{L_2(\mathbb{T})} + \varepsilon\|\nabla \jmath^\varepsilon\|^2_{L_2(\mathbb{T})}\big] = 0$.

(i) *Then the sequence* $\{\mathbb{P}^\varepsilon\} \subset \mathcal{P}(X)$ *satisfies a large deviations upper bound on* $(\mathcal{X}, d_{\mathcal{X}})$ *with speed* $\varepsilon^{-2\gamma+1}$ *and rate functional* $H : \mathcal{X} \to [0, +\infty]$

defined as

$$H(u) := \begin{cases} \displaystyle\int dv\, \varrho_u^+(v;dt,dx)\, \frac{D(v)}{a^2(v)} & \text{if } u \in \mathcal{E} \\ +\infty & \text{otherwise} \end{cases}$$

(ii) *Assume furthermore* $\lim_\varepsilon \varepsilon^{-3/2}\|\jmath^\varepsilon - 1\!\!1\|_{W^{-1,1}(\mathbb{T})} = 0$ *and* $f \in C^2([0,1])$. *Then the sequence* $\{\mathbb{P}^\varepsilon\} \subset \mathcal{P}(X)$ *satisfies a large deviations lower bound on* $(\mathcal{X}, d_{\mathcal{X}})$ *with speed* $\varepsilon^{-2\gamma+1}$ *and rate functional* $\overline{H} : \mathcal{X} \to [0,+\infty]$ *defined as*

$$\overline{H}(u) := \sup_{\substack{\mathcal{O} \ni u \\ \mathcal{O}\text{ open}}} \inf_{v \in \mathcal{O}\cap\mathcal{S}} H(v)$$

Since H is lower semicontinuous on $\mathcal{X}$, we have $\overline{H} \ge H$ on $\mathcal{X}$ and $\overline{H} = H$ on $\mathcal{S}$, namely a large deviations principle holds on $\mathcal{S}$. In order to obtain a full large deviations principle, one needs to show $H(u) \ge \overline{H}(u)$ for $u \notin \mathcal{S}$. This amounts to show that $\mathcal{S}$ is H-dense in $\mathcal{X}$, namely that for $u \in \mathcal{X}$ such that $H(u) < +\infty$ there exists a sequence $\{u^n\} \subset \mathcal{S}$ converging to u in $\mathcal{X}$ such that $H(u^n) \to H(u)$. In particular it can be shown that $\overline{H}(u) = H(u)$ for u piecewise smooth. The main difficulties here arise from the lacking of a chain rule formula connecting the measures ϱ_u to the structure of u itself. If u has bounded variation, Vol'pert chain rule [2] allows an explicit representation for ϱ_u and thus $H(u)$, see Remark 2.7 in [3]. On the other hand, there exists $u \in \mathcal{X}$ with infinite variation such that $H(u) < +\infty$, see Example 2.8 in [3]. While chain rule formulas out of the BV setting are subject to current research investigation, see e.g. [6, 1], only partial results are available.

Under the same hypotheses of Theorem 2.8, one can show that entropy-measure solutions to (2.3) are in $C([0,T]; L_1(\mathbb{T}))$, see Lemma 5.1 in [3]. By Kruzkov uniqueness theorem [5, Theorem 6.2.1], we gather that $H(u) = 0$ iff u is the Kruzkov solution to (2.3) with initial datum u_0. In particular, by item (i) in Theorem 2.8, large deviations principles with speeds slower than $\varepsilon^{-2\gamma+1}$ are trivial.

Note that in Proposition 2.1, Proposition 2.2, Theorem 2.4 and Theorem 2.8 various hypotheses on $\jmath^\varepsilon$ are required, the most restrictive in Theorem 2.8. It is easy to see that, if $\gamma > 1$, there exist convolution kernels $\jmath^\varepsilon$ satisfying them all.

3. Proofs

3.1. **Convergence and bounds.**

In the following we will need to consider several different perturbations of (2.2). In the next lemma we write down explicitly an Itô formula for (2.2). The corresponding Itô formula for the

perturbed equations can be obtained analogously, as the martingale term in these equations is always the same.

Lemma 3.1 (Itô formula). *Let $(\vartheta; Q)$ be an entropy sampler–entropy sampler flux pair for the equation* (2.3) *(recall in particular $\vartheta(u,T,x)=0$). Then*

$$\begin{aligned}
-\int dx\, \vartheta(u_0(x),0,x) - \int dt\, dx\, \big[\big(\partial_t \vartheta\big)\big(u^\varepsilon(t,x),t,x\big) + \big(\partial_x Q\big)\big(u^\varepsilon(t,x),t,x\big)\big] \\
= -\frac{\varepsilon}{2} \langle\langle \vartheta''(u^\varepsilon)\nabla u^\varepsilon, D(u^\varepsilon)\nabla u^\varepsilon\rangle\rangle - \frac{\varepsilon}{2}\langle\langle \partial_x \vartheta'(u^\varepsilon), D(u^\varepsilon)\nabla u^\varepsilon \rangle\rangle \\
+ \frac{\varepsilon^{2\gamma}}{2} \|\nabla \jmath^\varepsilon\|^2_{L_2(\mathbb{T})} \langle\langle \vartheta''(u^\varepsilon) a(u^\varepsilon), a(u^\varepsilon)\rangle\rangle \\
+ \frac{\varepsilon^{2\gamma}}{2} \|\jmath^\varepsilon\|^2_{L_2(\mathbb{T})} \langle\langle \vartheta''(u^\varepsilon)\nabla u^\varepsilon, [a'(u^\varepsilon)]^2 \nabla u^\varepsilon \rangle\rangle + N^{\varepsilon;\vartheta}(T)
\end{aligned} \tag{3.1}$$

where $N^{\varepsilon;\vartheta}$ is the martingale

$$N^{\varepsilon;\vartheta}(t) := -\varepsilon^\gamma \int_0^t \big\langle \jmath^\varepsilon * \big[a(u^\varepsilon)\vartheta''(u^\varepsilon)\nabla u^\varepsilon + a(u^\varepsilon)\partial_x\vartheta'(u^\varepsilon)\big], dW\big\rangle \tag{3.2}$$

Moreover the quadratic variation of $N^{\varepsilon,\vartheta}$ enjoys the bound

$$\big[N^{\varepsilon;\vartheta}, N^{\varepsilon;\vartheta}\big](t) \le \varepsilon^{2\gamma} \big\| a(u^\varepsilon)\big[\vartheta''(u^\varepsilon)\nabla u^\varepsilon + \partial_x \vartheta'(u^\varepsilon)\big]\big\|^2_{L_2([0,t]\times\mathbb{T})} \tag{3.3}$$

Proof. Equation (3.1) follows, up to minor manipulations, from Itô formula [7, Theorem 4.17] for the map

$$[0,T]\times U \ni (t,u) \mapsto \int dx\, \vartheta(u(x),t,x) \in \mathbb{R}$$

By (3.2) and (2.1), the quadratic variation of $N^{\varepsilon;\vartheta}$ is given by

$$\big[N^{\varepsilon;\vartheta}, N^{\varepsilon;\vartheta}\big](t) = \varepsilon^{2\gamma}\big\|\jmath^\varepsilon * \big\{a(u^\varepsilon)\big[\vartheta''(u^\varepsilon)\nabla u^\varepsilon + \partial_x\vartheta'(u^\varepsilon)\big]\big\}\big\|^2_{L_2([0,t]\times\mathbb{T})}$$

so that the inequality stated in the lemma follows by Young inequality for convolutions and hypothesis **H4)**. □

Lemma 3.2. *Let $\zeta, T > 0$, let X be a real, continuous, local, square integrable supermartingale starting from 0, and let $\tau \le T$ be a stopping time. Let $F : \mathbb{R} \to \mathbb{R}^+$ be such that:*

$$\frac{F(x)}{F(\zeta)} \le 2\frac{x}{\zeta} - 1, \quad \textit{for all } x > \zeta \tag{3.4}$$

Then:

$$\mathbb{P}\Big(\sup_{0\le t\le\tau} X(t) \ge \zeta,\ \big[X,X\big](\tau) \le F(\sup_{t\le\tau} X(t))\Big) \le \exp\Big[-\frac{\zeta^2}{2F(\zeta)}\Big] \tag{3.5}$$

Note that the hypotheses (3.4) on F are satisfied by any nonincreasing function, and by functions with affine or subaffine behavior. Lemma 3.2 provides an elementary generalization of the well known Bernstein inequality [19, page 153], which deals with the case of constant F.

Proof. Hypotheses on F imply that the map $G_\zeta : x \to \frac{\zeta}{F(\zeta)} x - \frac{1}{2}\frac{\zeta^2}{F(\zeta)^2} F(x)$ satisfies $G_\zeta(x) \geq G_\zeta(\zeta) = \frac{\zeta^2}{2F(\zeta)}$ for all $x \geq \zeta$. Therefore:

$$\begin{aligned}
&\mathbb{P}\Big(\sup_{t\leq\tau} X(t) \geq \zeta,\ \big[X,X\big](\tau) \leq F(\sup_{t\leq\tau} X(t))\Big) \\
&\leq \mathbb{P}\Big(e^{\frac{\zeta}{F(\zeta)}\sup_{t\leq\tau} X(t) - \frac{1}{2}\frac{\zeta^2}{F(\zeta)^2}F(\sup_{t\leq\tau} X(t))} \geq e^{\frac{1}{2}\frac{\zeta^2}{F(\zeta)}}, \\
&\qquad \big[X,X\big](\tau) \leq F(\sup_{t\leq\tau} X(t))\Big) \\
&\leq \mathbb{P}\Big(\sup_{t\leq T} e^{\frac{\zeta}{F(\zeta)}X(t) - \frac{1}{2}\frac{\zeta^2}{F(\zeta)^2}[X,X](t)} \geq e^{\frac{1}{2}\frac{\zeta^2}{F(\zeta)}}\Big) \leq e^{-\frac{\zeta^2}{2F(\zeta)}}.
\end{aligned}$$

where in the last line we applied the maximal inequality for positive supermartingales [19, page 58], to the supermartingale $e^{\frac{\zeta}{F(\zeta)}X(t) - \frac{1}{2}\frac{\zeta^2}{F(\zeta)^2}[X,X](t)}$. □

The next lemma provides a key a priori bound.

Lemma 3.3. *For $\varepsilon > 0$, let $E^\varepsilon \in L_2\big([0,T]; H^1(\mathbb{T})\big)$ and let $\mathbb{Q}^\varepsilon \in \mathcal{P}\big(C\big([0,T]; U\big)\big)$ be any martingale solution to the Cauchy problem*

$$\begin{aligned}
du &= \big[-\nabla\cdot f(u) + \frac{\varepsilon}{2}\nabla\cdot\big(D(u)\nabla u\big) - \nabla\cdot(a(u)E^\varepsilon)\big]\,dt \\
&\quad + \varepsilon^\gamma\,\nabla\cdot\big[a(u)(\jmath^\varepsilon * dW)\big] \\
u(0,x) &= u_0^\varepsilon(x)
\end{aligned} \tag{3.6}$$

Assume $\|\nabla E^\varepsilon\|_{L_2([0,T]\times\mathbb{T})} \leq C_0$ for some constant C_0 independent of ε, and $\lim_\varepsilon \varepsilon^{2\gamma-1}(\|\jmath^\varepsilon\|^2_{L_2(\mathbb{T})} + \varepsilon\|\nabla\jmath^\varepsilon\|^2_{L_2(\mathbb{T})}) = 0$. Then there exist C, $\varepsilon_0 > 0$ such that for any $\varepsilon < \varepsilon_0$:

$$\varepsilon\|\nabla u\|^2_{L_2([0,T]\times\mathbb{T})} \leq C + N^\varepsilon(T,u) \qquad \textit{for } \mathbb{Q}^\varepsilon \textit{ a.e. } u \tag{3.7}$$

where N^ε is a $\mathbb{Q}^\varepsilon$-martingale starting from 0 and satisfying

$$\mathbb{Q}^\varepsilon\big(\sup_{t\leq T} N^\varepsilon(t) > \zeta\big) \leq \exp\Big\{ -\frac{\zeta^2}{\varepsilon^{2\gamma-1}C(1+\zeta)}\Big\} \tag{3.8}$$

Proof. Itô formula for the map $U \ni u \mapsto \int dx\, u^2(x) \in \mathbb{R}$ can be obtained as in Lemma 3.1, so that

$$\begin{aligned}
&\|u(T)\|_{L_2(\mathbb{T})} - \|u_0\|_{L_2(\mathbb{T})} + \varepsilon\langle\langle \nabla u, D(u)\nabla u\rangle\rangle \\
&\quad = -\langle\langle A(u), \nabla E^\varepsilon\rangle\rangle + \varepsilon^{2\gamma}\|\nabla\jmath^\varepsilon\|^2_{L_2(\mathbb{T})}\,\|a(u)\|^2_{L_2([0,T]\times\mathbb{T})} \\
&\quad\quad + \varepsilon^{2\gamma}\|\jmath^\varepsilon\|^2_{L_2(\mathbb{T})}\,\|a'(u)\nabla u\|^2_{L_2([0,T]\times\mathbb{T})} + N^\varepsilon(T,u)
\end{aligned}$$

where $A \in C^1([0,1])$ is any antiderivative of $a(\cdot)$ and N^ε is a $\mathbb{Q}^\varepsilon$-martingale, which -reasoning as in the proof of (3.3)- satisfies

$$\big[N^\varepsilon, N^\varepsilon\big](T,u) \le 4\,\varepsilon^{2\gamma} \|a(u)\nabla u\|^2_{L_2([0,T]\times\mathbb{T})} \tag{3.9}$$

By **H2)**, **H3)** and the hypotheses of this lemma, there exist C_1, $\varepsilon_0 > 0$ such that, for each $\varepsilon \le \varepsilon_0$ and $v \in [0,1]$

$$\varepsilon^{2\gamma} \|\jmath^\varepsilon\|^2_{L_2(\mathbb{T})} [a'(v)]^2 \le \frac{\varepsilon}{2} D(v)$$

$$|\langle\langle A(u), \nabla E^\varepsilon \rangle\rangle| + \varepsilon^{2\gamma} \|\nabla \jmath^\varepsilon\|^2_{L_2(\mathbb{T})}\, \|a(u)\|^2_{L_2([0,T]\times\mathbb{T})} \le C_1$$

Therefore, since $|u_0| \le 1$

$$\frac{\varepsilon}{2} \langle\langle \nabla u, D(u)\nabla u \rangle\rangle \le 1 + C_1 + N^\varepsilon(T,u) \tag{3.10}$$

and thus (3.7) since D is uniformly positive. By (3.9) and (3.10), there exists a constant $C_2 > 0$ such that

$$\big[N^\varepsilon, N^\varepsilon\big](T,u) \le \varepsilon^{2\gamma} C_2 \langle\langle \nabla u, D(u)\nabla u \rangle\rangle \le 2\, C_2\, \varepsilon^{2\gamma-1} \big[1 + C_1 + N^\varepsilon(T)\big] \tag{3.11}$$

This inequality allows the application of Lemma 3.2 for the martingale N^ε with

$$F(\zeta) = 2\, C_2\, \varepsilon^{2\gamma-1} (1 + C_1 + \zeta)$$

which clearly satisfies the condition (3.4). The bound (3.8) then follows straightforwardly. □

The following lemma provides a stability result for (2.2). It will be repeatedly used to evaluate the effects of the Girsanov terms appearing in (2.2) when absolutely continuous perturbations of $\mathbb{P}^\varepsilon$ are considered.

Lemma 3.4. *For each $\varepsilon > 0$, let $v^\varepsilon : \mathcal{X} \to \mathcal{X} \cap L_2([0,T]; H^1(\mathbb{T}))$ and $G^\varepsilon : \mathcal{X} \times \mathcal{X} \to L_2([0,T]\times\mathbb{T})$ be adapted maps (with respect to the standard filtrations of $\mathcal{X}$ and $\mathcal{X}\times\mathcal{X}$ respectively). Let $\mathbb{Q}^\varepsilon \in \mathcal{P}(\mathcal{X})$ be any martingale solution to the stochastic Cauchy problem in the unknown u*

$$\begin{aligned} du = \Big[&- \nabla\cdot f(u) + \frac{\varepsilon}{2} \nabla \cdot \big(D(u)\nabla u\big) + \partial_t v^\varepsilon(u) + \nabla \cdot f(v^\varepsilon(u)) \\ &- \frac{\varepsilon}{2} \nabla \cdot \big(D(v^\varepsilon(u)) \nabla v^\varepsilon(u)\big) + G^\varepsilon(u, v^\varepsilon(u))\Big]\, dt \\ &+ \varepsilon^\gamma\, \nabla \cdot \big[a(u)(\jmath^\varepsilon * dW)\big] \\ u(0,x) = {} & u_0^\varepsilon(x) \end{aligned} \tag{3.12}$$

Suppose

(i) $\lim_\varepsilon \varepsilon^{2(\gamma-1)} \big[\|\jmath^\varepsilon\|^2_{L_2} + \varepsilon \|\nabla \jmath^\varepsilon\|^2_{L_2}\big] = 0.$

(ii) *There exist adapted processes* G_1^ε, G_2^ε, $G_3^\varepsilon : \chi \times \chi \to L_2([0,T]\times\mathbb{T})$ *such that* $G^\varepsilon(u,v)(t,x) = G_1^\varepsilon(u,v)(t,x) + \nabla\cdot G_2^\varepsilon(u,v)(t,x) + \nabla\cdot G_3^\varepsilon(u,v)(t,x)$, *and*

$$|G_3^\varepsilon(u,v^\varepsilon(u))(t,x)| \le G_4^\varepsilon(u)(t,x)|u - v^\varepsilon(u)| \qquad \textit{for } \mathbb{Q}^\varepsilon \textit{ a.e. } u$$

for some adapted process $G_4^\varepsilon : \mathcal{X} \to L_2([0,T]\times\mathbb{T})$.

(iii) *Let* G_1, G_2 *be as in (ii). Then for each* $\delta > 0$

$$\begin{aligned}\lim_{\varepsilon} \mathbb{Q}^\varepsilon\big(&\|v^\varepsilon(u)(0) - u_0^\varepsilon\|_{L_1(\mathbb{T})} + \|G_1^\varepsilon(u,v^\varepsilon(u))\|_{L_1([0,T]\times\mathbb{T})} \\ &+ \varepsilon^{-1}\|G_2^\varepsilon(u,v^\varepsilon(u))\|_{L_2([0,T]\times\mathbb{T})} > \delta\big) = 0\end{aligned}$$

(iv) *Let* G_4 *be as in (ii). Then*

$$\lim_{\ell\to+\infty}\overline{\lim_{\varepsilon}}\,\mathbb{Q}^\varepsilon\big(\|G_4^\varepsilon(u,v^\varepsilon(u))\|_{L_2([0,T]\times\mathbb{T})} + \varepsilon\|\nabla u\|_{L_2([0,T]\times\mathbb{T})} > \ell\big) = 0$$

Then for each $\delta > 0$

$$\lim_{\varepsilon}\mathbb{Q}^\varepsilon\big(\|u - v^\varepsilon(u)\|_{L_\infty([0,T];L_1(\mathbb{T}))} > \delta\big) = 0 \tag{3.13}$$

Proof. We denote by $z^\varepsilon(t,x) \equiv z^\varepsilon(u)(t,x) := u(t,x) - v^\varepsilon(u)(t,x) \in [-1,1]$. Let $l \in C^2([-1,1])$. For each ε, $t > 0$ let us define (in the following we omit the dependence of v^ε and z^ε on the u variable)

$$\begin{aligned}N^{\varepsilon;l}(t,u) \;:=\;& \int dx\,[l(z^\varepsilon(t)) - l(z^\varepsilon(0))] - \int_0^t ds\Big[\langle l''(z^\varepsilon)\nabla z^\varepsilon, f(u) - f(v^\varepsilon)\rangle \\ & -\frac{\varepsilon}{2}\langle l''(z^\varepsilon)\nabla z^\varepsilon, D(v^\varepsilon)\nabla z^\varepsilon\rangle - \frac{\varepsilon}{2}\langle l''(z^\varepsilon)\nabla z^\varepsilon, [D(u) - D(v^\varepsilon)]\nabla u^\varepsilon\rangle \\ & +\langle l'(z^\varepsilon), G_1^\varepsilon(u,v^\varepsilon)\rangle - \langle l''(z^\varepsilon)\nabla z^\varepsilon, G_2^\varepsilon(u,v^\varepsilon)\rangle \\ & -\langle l''(z^\varepsilon)\nabla z^\varepsilon, G_3^\varepsilon(u,v^\varepsilon)\rangle + \frac{\varepsilon^{2\gamma}}{2}\|\nabla \jmath^\varepsilon\|^2_{L_2(\mathbb{T})}\langle l''(z^\varepsilon)a(u), a(u)\rangle \\ & +\frac{\varepsilon^{2\gamma}}{2}\|\jmath^\varepsilon\|^2_{L_2(\mathbb{T})}\langle l''(z^\varepsilon)\nabla u, [a'(u)]^2\nabla u\rangle\Big]\end{aligned}$$

By Itô formula, $N^{\varepsilon;l}$ is a $\mathbb{Q}^\varepsilon$-martingale starting at 0, and applying Young inequality for convolutions (analogously to (3.3))

$$\big[N^{\varepsilon,l}, N^{\varepsilon,l}\big](t,u) \le \varepsilon^{2\gamma}\|a(u)l''(z^\varepsilon)\nabla z^\varepsilon\|^2_{L_2([0,t]\times\mathbb{T})} \tag{3.14}$$

We now choose l convex and define

$$R^{\varepsilon,l}(t) \equiv R^{\varepsilon,l}(u)(t) := \Big[\int dx\,\langle l''(z^\varepsilon(t))\nabla z^\varepsilon(t), \nabla z^\varepsilon(t)\rangle\Big]^{1/2}$$

Since D and f are Lipschitz, and D is uniformly positive, by (3.14) and Cauchy-Schwartz inequality we gather

$$\begin{aligned}\int dx\, l(z^\varepsilon(t)) - l(z^\varepsilon(0)) \le & -c\,\varepsilon\,[R^{\varepsilon;l}(t)]^2 \|\sqrt{l''(z^\varepsilon)}z^\varepsilon\|_{L_\infty([0,T]\times\mathbb{T})} R^{\varepsilon;l}(t)\\ & + C_1\varepsilon\|\nabla u\|_{L_2([0,T]\times\mathbb{T})}\,\|\sqrt{l''(z^\varepsilon)}z^\varepsilon\|_{L_\infty([0,T]\times\mathbb{T})}R^{\varepsilon;l}(t)\\ & + \|l'(z^\varepsilon)\|_{L_\infty([0,T]\times\mathbb{T})}\,\|G_1^\varepsilon(u,v^\varepsilon)\|_{L_1([0,T]\times\mathbb{T})}\\ & + \|G_2^\varepsilon(u,v^\varepsilon)\|_{L_2([0,T]\times\mathbb{T})}\|\sqrt{l''(z^\varepsilon)}\|_{L_\infty([0,T]\times\mathbb{T})}R^{\varepsilon;l}(t)\\ & + \|G_4^\varepsilon(u)\|_{L_2([0,T]\times\mathbb{R})}\,\|\sqrt{l''(z^\varepsilon)}z^\varepsilon\|_{L_\infty([0,T]\times\mathbb{T})}R^{\varepsilon;l}(t)\\ & + C_1\varepsilon^{2\gamma}\|\nabla \jmath^\varepsilon\|^2_{L_2(\mathbb{T})}\|l''(z^\varepsilon)\|_{L_\infty([0,T]\times\mathbb{T})}\\ & + C_1\varepsilon^{2\gamma}\|\jmath^\varepsilon\|^2_{L_2(\mathbb{T})}\|l''(z^\varepsilon)\|_{L_\infty([0,T]\times\mathbb{T})}\|\nabla u\|^2_{L_2([0,T]\times\mathbb{T})} + N^{\varepsilon;l}(t)\end{aligned} \tag{3.15}$$

for some constants c, $C_1 > 0$ independent of ε and l. For arbitrary $\zeta > 0$ to be chosen below, we now consider $l(Z) = \sqrt{Z^2 + \varepsilon^2\zeta^2}$ so that

$$\begin{aligned} & |Z| \le l(Z) \le |Z| + \varepsilon\zeta & & \max_{Z\in[-1,1]} |l'(Z)| \le 1 \\ & \max_{Z\in[-1,1]} |l''(Z)| \le \varepsilon^{-1}\zeta^{-1} & & \max_{Z\in[-1,1]} |l''(Z)\,Z^2| \le \sqrt{2}\varepsilon\zeta \end{aligned}$$

Using these bounds in the right hand side of (3.15), we get for some $C_2 > 0$

$$\begin{aligned}\int dx\,|z^\varepsilon(t)| \le & \int dx\,|z^\varepsilon(0)| + C_2\|G_1^\varepsilon\|_{L_1([0,T]\times\mathbb{T})}\\ & + C_2\big[1 + \varepsilon^2\|\nabla u\|^2_{L_2([0,T]\times\mathbb{T})} + \|G_4^\varepsilon(u)\|^2_{L_2([0,T]\times\mathbb{T})}\big]\zeta\\ & + C_2\zeta^{-1}\big[\varepsilon^{-2}\|G_2^\varepsilon\|^2_{L_2([0,T]\times\mathbb{T})} + \varepsilon^{2\gamma-1}\|\nabla\jmath^\varepsilon\|^2_{L_2(\mathbb{T})}\\ & \qquad + \varepsilon^{2(\gamma-1)}\|\jmath^\varepsilon\|^2_{L_2(\mathbb{T})}\|\nabla u\|^2_{L_2([0,T]\times\mathbb{T})}\big]\\ & - \frac{c\,\varepsilon}{2}[R^{\varepsilon;l}(t)]^2 + N^{\varepsilon;l}(t)\end{aligned} \tag{3.16}$$

where we have also used the straightforward inequality $\alpha R - \frac{c\varepsilon}{2}R^2 \le \frac{\alpha}{2c\varepsilon}$ for a suitable $\alpha \in \mathbb{R}$.

Recalling (3.14), for some $C_3 > 0$ independent of ε, ζ

$$\big[N^{\varepsilon;l}, N^{\varepsilon;l}\big](t,u) \le C_3\,\varepsilon^{2\gamma}\zeta^{-1}[R^{\varepsilon;l}(t)]^2$$

so that, by maximal inequality for positive supermartingales [19, page 58], for each $\delta > 0$ the term in the last line of (3.16) satisfies

$$\begin{aligned}
&\mathbb{Q}^\varepsilon\Big(\sup_{s\le t} N^{\varepsilon;l}(s) - \frac{c\,\varepsilon}{2}[R^{\varepsilon;l}(s)]^2 > \delta\Big) \le \\
&\mathbb{Q}^\varepsilon\bigg(\sup_{s\le t}\exp\Big(\frac{2\,c}{C_3}\varepsilon^{1-2\gamma}\zeta\,N(s) - \frac{2\,c^2}{C_3^2}\varepsilon^{2(1-2\gamma)}\zeta^2\,[N,N](s)\Big) > \\
&\qquad \exp(\frac{2\,c}{C_3}\varepsilon^{1-2\gamma}\zeta\,\delta)\bigg) \le \exp(-\frac{2\,c}{C_3}\varepsilon^{-2\gamma+1}\zeta\,\delta)
\end{aligned} \tag{3.17}$$

Furthermore for $\ell > 0$

$$\begin{aligned}
\mathbb{Q}^\varepsilon\Big(\sup_t \int dx\,|z^\varepsilon(t)| > \delta\Big) \le \mathbb{Q}^\varepsilon\Big(\sup_t \int dx\,|z^\varepsilon(t)| > \delta,& \\
\|G_4^\varepsilon(u, v^\varepsilon(u))\|_{L_2([0,T]\times\mathbb{T})} + \varepsilon\|\nabla u\|_{L_2([0,T]\times\mathbb{T})} \le \ell\Big) &+ o_{\ell,\varepsilon}
\end{aligned}$$

where $\lim_\ell \overline{\lim}_\varepsilon\, o_{\ell,\varepsilon} = 0$ by hypotheses (iv). Therefore, using hypotheses (i) and (iii) and the estimate (3.17) in (3.16), the result easily follows as we let $\varepsilon \to 0$, then $\zeta \to 0$ and finally $\ell \to +\infty$. □

The following result will be used to provide exponential tightness in stronger topologies in the next sections.

Lemma 3.5. *There exists a sequence $\{K_\ell\}$ of compact subsets of $C\big([0,T];U\big)$ such that*

$$\lim_\ell \overline{\lim_\varepsilon}\, \varepsilon^{2\gamma} \log \mathbb{P}^\varepsilon(K_\ell^c) = -\infty$$

Proof. We refer to the criterion in [9, Corollary 4.17] to establish the exponential tightness of $\{\mathbb{P}^\varepsilon\}$. Let $d \in C^1([0,1])$ be any antiderivative of D. Integrating twice by parts the diffusive term in the weak formulation of (2.2) (see (A.2) and (A.3)), for each $\varphi \in C^\infty(\mathbb{T})$ the map $E^{\varepsilon;\varphi} : [0,T] \times C\big([0,T];U\big) \to \mathbb{R}$ defined by

$$\begin{aligned}
E^{\varepsilon;\varphi}(t;u) := \exp\Big[&\varepsilon^{-2\gamma}\langle u(t),\varphi\rangle - \varepsilon^{-2\gamma}\langle u(0),\varphi\rangle \\
&-\varepsilon^{-2\gamma}\int_0^t ds\,\langle f(u) + \frac{\varepsilon}{2}d(u), \Delta\varphi\rangle - \frac{1}{2}\langle \jmath * (a(u)\nabla\varphi), \jmath * (a(u)\nabla\varphi)\rangle\Big]
\end{aligned}$$

is a martingale. For a fixed $\varphi \in C^\infty(\mathbb{T})$, the following bound on the integral term in the definition of $E^{\varepsilon;\varphi}$ is easily established

$$\sup_{v\in U}\Big|\langle f(v) + \frac{\varepsilon}{2}d(v), \Delta\varphi\rangle - \frac{1}{2}\langle \jmath * (a(v)\nabla\varphi), \jmath * (a(v)\nabla\varphi)\rangle\Big| < +\infty$$

Furthermore the family of maps $l^\varphi : U \ni v \to \langle v,\varphi\rangle \in \mathbb{R}$ is closed under addition, separates points in U and satisfies $c\,l^\varphi = l^{c\varphi}$ for $c \in \mathbb{R}$. All the hypotheses of the criterion in [9, Corollary 4.17] are therefore satisfied. □

Proof of Proposition 2.2. with $\mathbb{Q}^\varepsilon \equiv \mathbb{P}^\varepsilon := P \circ (u^\varepsilon)^{-1}$, and v^ε as the solution to the (deterministic) Cauchy problem

$$\begin{aligned} &\partial_t v = -\nabla \cdot f(v) + \frac{\varepsilon}{2} \nabla \cdot \big(D(v) \nabla v\big) \\ &v(0,x) = u_0(x) \end{aligned}$$

$\mathbb{P}^\varepsilon$ and v^ε fulfill the hypotheses Lemma 3.4, since $G^\varepsilon \equiv 0$ and Lemma 3.3 holds (with $E^\varepsilon \equiv 0$). As well known [5, Chap. 6.3], $v^\varepsilon \to \bar{u}$ in $L_p([0,T] \times \mathbb{T})$. Therefore the statement of the proposition follows by the same Lemma 3.4 and the fact that $\mathbb{P}^\varepsilon$ is (exponentially) tight in $C\big([0,T]; U\big)$, as proved in Lemma 3.5. □

3.2. **Large deviations with speed $\varepsilon^{-2\gamma}$.** In this section we prove Theorem 2.4.

Lemma 3.6. *There exists a sequence $\{\mathcal{K}_\ell\}$ of compact subsets of $\mathcal{M}$ such that*

$$\lim_\ell \overline{\lim_\varepsilon}\, \varepsilon^{2\gamma} \log \mathbf{P}^\varepsilon(\mathcal{K}_\ell^c) = -\infty \tag{3.18}$$

Proof. Let the sequence $\{K_\ell\}$ of compact subsets of $C\big([0,T]; U\big)$ be as in Lemma 3.5. For $\ell > 0$ consider the set

$$\tilde{\mathcal{K}}_\ell := \{\mu \in \mathcal{M} \,:\, \mu_{t,x} = \delta_{u(t,x)} \text{ for some } u \in K_\ell\}$$

Then $\mathbf{P}^\varepsilon(\tilde{\mathcal{K}}_\ell) = \mathbb{P}^\varepsilon(K_\ell)$ and by Lemma 3.5

$$\lim_\ell \overline{\lim_\varepsilon}\, \varepsilon^{2\gamma} \log \mathbf{P}^\varepsilon(\tilde{\mathcal{K}}_\ell^c) = -\infty$$

On the other hand $\tilde{\mathcal{K}}_\ell$ is precompact in $(\mathcal{M}, d_\mathcal{M})$ for any ℓ, and thus the Lemma is proved by taking $\mathcal{K}_\ell$ to be the closure of $\tilde{\mathcal{K}}_\ell$. □

Proof of Theorem 2.4: upper bound. Let $d \in C^2([0,1])$ be any antiderivative of D. For $\varepsilon > 0$ and $\varphi \in C^\infty([0,T] \times \mathbb{T})$, define the map $\mathcal{N}^{\varepsilon;\varphi} : [0,T] \times \mathcal{M} \to \mathbb{R}$ by

$$\begin{aligned} \mathcal{N}^{\varepsilon;\varphi}(t,\mu) \;:=\; &\langle \mu_{T,\cdot}(\imath), \varphi(T)\rangle - \langle u_0, \varphi(0)\rangle \\ &- \int_0^t ds\, \Big[\langle \mu(\imath), \partial_t \varphi\rangle - \langle \mu(f), \nabla\varphi\rangle + \frac{\varepsilon}{2}\langle \mu(d), \Delta\varphi\rangle\Big] \end{aligned}$$

$\mathbf{P}^\varepsilon$ is concentrated on the set

$$\{\mu \in \mathcal{M} \,:\, \mu = \delta_u \text{ for some } u \in C\big([0,T]; U\big)\}$$

so that $\mathcal{N}^{\varepsilon;\varphi}$ is a $\mathbf{P}^\varepsilon$-martingale. Indeed an integration by parts shows that $\mathcal{N}^{\varepsilon;\varphi}(t, \delta_u)$ is the martingale term appearing in the very definition of *martingale solution to* (2.2), see the appendix. Reasoning as in (3.3), we have

$$\big[\mathcal{N}^{\varepsilon;\varphi}, \mathcal{N}^{\varepsilon;\varphi}\big](t,\mu) \le \varepsilon^{2\gamma} \int_0^t ds\, \langle \mu(a^2) \nabla\varphi, \nabla\varphi\rangle$$

Therefore, the map $\mathcal{Q}^{\varepsilon;\varphi} : [0,T] \times \mathcal{M} \to \mathbb{R}$ defined by

$$\mathcal{Q}^{\varepsilon;\varphi}(t,\mu) := \exp\Big\{\mathcal{N}^{\varepsilon;\varphi}(t,\mu) - \frac{\varepsilon^{2\gamma}}{2}\int_0^t ds\, \langle \mu(a^2)\nabla\varphi, \nabla\varphi\rangle\Big\}$$

is a continuous $\mathbf{P}^\varepsilon$-supermartingale, with $\mathcal{Q}^{\varepsilon;\varphi}(0,\mu) = 1$ and $\mathcal{Q}^{\varepsilon;\varphi}(T,\mu) > 0$, $\mathbf{P}^\varepsilon$ a.s.. For an arbitrary Borel set $\mathcal{A} \subset \mathcal{M}$ we then have

$$\begin{aligned}
\mathbf{P}^\varepsilon(\mathcal{A}) &= \mathbb{E}^{\mathbf{P}^\varepsilon}\big(1\!\!1_{\mathcal{A}}(\cdot)\mathcal{Q}^{\varepsilon;\varphi}(T,\cdot)[\mathcal{Q}^{\varepsilon;\varphi}(T,\cdot)]^{-1}\big) \\
&\le \sup_{\mu\in\mathcal{A}}[\mathcal{Q}^{\varepsilon;\varphi}(T,\mu)]^{-1}\mathbb{E}^{\mathbf{P}^\varepsilon}\big(1\!\!1_{\mathcal{A}}(\cdot)\mathcal{Q}^{\varepsilon;\varphi}(T,\cdot)\big) \le \sup_{\mu\in\mathcal{A}}[\mathcal{Q}^{\varepsilon;\varphi}(T,\mu)]^{-1}
\end{aligned}$$

Since this inequality holds for each φ, we can evaluate it replacing φ with $\varepsilon^{-2\gamma}\varphi$, thus obtaining

$$\begin{aligned}
\varepsilon^{2\gamma}\log \mathbf{P}^\varepsilon(\mathcal{A}) &\le -\inf_{\mu\in\mathcal{A}}\Big\{\langle \mu_{T,\cdot}(\imath), \varphi(T)\rangle - \langle u_0, \varphi(0)\rangle - \langle\langle \mu(\imath), \partial_t\varphi\rangle\rangle \\
&\qquad -\langle\langle \mu(f), \nabla\varphi\rangle\rangle - \frac{\varepsilon}{2}\langle\langle \mu(d), \Delta\varphi\rangle\rangle - \frac{1}{2}\langle\langle \mu(a^2)\nabla\varphi, \nabla\varphi\rangle\rangle\Big\} \\
&\le -\inf_{\mu\in\mathcal{A}}\Big\{\langle \mu_{T,\cdot}(\imath), \varphi(T)\rangle - \langle u_0, \varphi(0)\rangle - \langle\langle \mu(\imath), \partial_t\varphi\rangle\rangle \\
&\qquad -\langle\langle \mu(f), \nabla\varphi\rangle\rangle - \frac{1}{2}\langle\langle \mu(a^2)\nabla\varphi, \nabla\varphi\rangle\rangle\Big\} + \varepsilon\, C_{d,\varphi}
\end{aligned}$$

for some constant $C_{d,\varphi}$ depending only on d and φ. Taking the limsup for $\varepsilon \to 0$, the last term vanishes. Optimizing on φ:

$$\begin{aligned}
\overline{\lim_{\varepsilon}}\, \varepsilon^{2\gamma}\log \mathbf{P}^\varepsilon(\mathcal{A}) &\le -\sup_{\varphi\in C^\infty([0,T]\times\mathbb{T})}\inf_{\mu\in\mathcal{A}}\Big\{\langle \mu_{T,\cdot}(\imath), \varphi(T)\rangle - \langle u_0, \varphi(0)\rangle \\
&\qquad -\langle\langle \mu(\imath), \partial_t\varphi\rangle\rangle - \langle\langle \mu(f), \nabla\varphi\rangle\rangle - \frac{1}{2}\langle\langle \mu(a^2)\nabla\varphi, \nabla\varphi\rangle\rangle\Big\}
\end{aligned}$$

By a standard application [13, Appendix 2, Lemma 3.2] of the minimax lemma, we gather that upper bound with rate $\mathcal{I}$, see (2.6), holds on each compact subset $\mathcal{K} \subset \mathcal{M}$. By Lemma 3.6, it holds on each closed subset of $\mathcal{M}$. □

We recall a well known method to prove large deviations lower bounds, see e.g. [12, Chap. 4]. For $\mathbb{P}$, $\mathbb{Q}$ two Borel probability measures on a Polish space, we denote by $\mathrm{Ent}(\mathbb{Q}|\mathbb{P})$ the relative entropy of $\mathbb{Q}$ with respect to $\mathbb{P}$.

Lemma 3.7. *Let $\mathcal{X}$ be a Polish space, $I : \mathcal{X} \to [0,+\infty]$ a positive functional, $\{\alpha_\varepsilon\}$ a sequence of positive reals such that $\lim_\varepsilon \alpha_\varepsilon = 0$, and let $\{\mathbb{P}^\varepsilon\} \subset \mathcal{P}(\mathcal{X})$. Suppose that for each $x \in \mathcal{X}$ there is a sequence $\{\mathbb{Q}^{\varepsilon,x}\} \subset \mathcal{P}(\mathcal{X})$ such that $\mathbb{Q}^{\varepsilon,x} \to \delta_x$ weakly in $\mathcal{P}(\mathcal{X})$, and $\overline{\lim}_\varepsilon\, \alpha_\varepsilon \mathrm{Ent}_\varepsilon(\mathbb{Q}^{\varepsilon,x}|\mathbb{P}^\varepsilon) \le I(x)$. Then $\{\mathbb{P}^\varepsilon\}$ satisfies a large deviations lower bound with speed α_ε^{-1} and rate I.*

Proof of Theorem 2.4: lower bound. We will prove the lower bound following the strategy suggested by Lemma 3.7. More precisely, consider the set

$$\mathcal{M}_0 := \Big\{\mu \in \mathcal{M} \,:\, \exists \zeta > 0 \,:\, \mu = \delta_u \text{ for some } u \in C^2\big([0,T]\times\mathbb{T};[\zeta,1-\zeta]\big)\Big\}$$

Here we prove that for each $\mu \in \mathcal{M}_0$ there exists a sequence of probability measures $\{\mathbf{Q}^\varepsilon\} \subset \mathcal{P}(\mathcal{M})$ such that $\mathbf{Q}^\varepsilon \to \delta_\mu$ and $\overline{\lim}\, \varepsilon^{2\gamma}\mathrm{Ent}(\mathbf{Q}^\varepsilon|\mathbf{P}^\varepsilon) \le \mathcal{I}(\mu)$. By Lemma 3.7 this will yield a large deviations lower bound with rate $\tilde{\mathcal{I}} : \mathcal{M} \to [0,+\infty]$ defined as

$$\tilde{\mathcal{I}}(\mu) := \begin{cases} \mathcal{I}(\mu) & \text{if } \mu \in \mathcal{M}_0 \\ +\infty & \text{otherwise} \end{cases}$$

By a standard diagonal argument, the lower bound then also holds with the lower semicontinuous envelope of $\tilde{\mathcal{I}}$ as rate functional. In [3, Theorem 4.1] it is shown, in a slightly different setting, that the lower semicontinuous envelope of $\tilde{\mathcal{I}}$ is indeed $\mathcal{I}$. By the assumption $\zeta \le u_0 \le 1-\zeta$ (which is equivalent to the requirement that $a^2(u_0)$ is uniformly positive), it is not difficult to adapt the arguments in the proof of Theorem 4.1 in [3, Theorem 4.1], to obtain the analogous result in this case. We are thus left with the proof of the lower bound on $\mathcal{M}_0$.

Let $\mu \in \mathcal{M}_0$ be such that $\mathcal{I}(\mu) < \infty$. Then $\mu = \delta_v$ for some smooth $v \in C\big([0,T];U\big)$ with $v(0,x) = u_0(x)$ and $a(v)^2 \ge r$ for some $r > 0$. By the definition of $\mathcal{I}$ and the smoothness of v

$$\begin{aligned} \mathcal{I}(\mu) &= \sup_{\varphi \in C^\infty\left([0,T]\times\mathbb{T}\right)} \Big\{ -\langle\langle \partial_t v + \nabla\cdot f(v), \varphi\rangle\rangle - \frac{1}{2}\langle\langle a(v)^2\nabla\varphi, \nabla\varphi\rangle\rangle \Big\} \\ &\ge \sup_{\varphi \in C^\infty\left([0,T]\times\mathbb{T}\right)} \Big\{ -\langle\langle \partial_t v + \nabla\cdot f(v), \varphi\rangle\rangle - \frac{r}{2}\langle\langle \nabla\varphi, \nabla\varphi\rangle\rangle \Big\} \end{aligned}$$

Since the supremums in this formula are finite, Riesz representation lemma implies the existence of a $\Psi^v \in L_2\big([0,T];H^1(\mathbb{T})\big)$ such that

$$\partial_t v + \nabla\cdot f(v) = -\nabla\cdot[a(v)^2\nabla\Psi^v] \tag{3.19}$$

holds weakly and

$$\mathcal{I}(\mu) = \frac{1}{2}\langle\langle a(v)^2\nabla\Psi^v, \nabla\Psi^v\rangle\rangle \tag{3.20}$$

We next define the P-martingale $M^{\varepsilon;v}$ on Ω as

$$M^{\varepsilon;v}(t) := -\varepsilon^{-\gamma}\int_0^t \big\langle \jmath^\varepsilon * [a(v)\nabla\Psi^v], dW\big]\big\rangle$$

so that, by Young inequality for convolutions and (3.20), we have P a.s.

$$\big[M^{\varepsilon;v}, M^{\varepsilon;v}\big](T) \le \varepsilon^{-2\gamma}\|a(v)\nabla\Psi^v\|^2_{L_2([0,T]\times\mathbb{T})} = 2\varepsilon^{-2\gamma}\mathcal{I}(\mu) \tag{3.21}$$

Since the quadratic variation of $M^{\varepsilon;v}$ is bounded, its stochastic exponential

$$E^{\varepsilon;v}(t,\omega) := \exp\big(M^{\varepsilon;v}(t,\omega) - \frac{1}{2}[M^{\varepsilon;v}, M^{\varepsilon;v}](t,\omega)\big)$$

is a uniformly integrable P-martingale. For $\varepsilon > 0$ we define the probability measure $Q^{\varepsilon;v}$ on Ω by

$$Q^{\varepsilon;v}(d\omega) := E^{\varepsilon;v}(T,\omega)P(d\omega)$$

Recalling that u^ε was the process solving (2.2), we next define $\mathbf{Q}^{\varepsilon;v} := Q^{\varepsilon;v} \circ (\delta_{u^\varepsilon})^{-1} \in \mathcal{P}(\mathcal{M})$. Then

$$\begin{aligned}
\varepsilon^{2\gamma}\mathrm{Ent}(\mathbf{Q}^{\varepsilon;v}|\mathbf{P}^{\varepsilon;v}) &\le \varepsilon^{2\gamma}\mathrm{Ent}(Q^{\varepsilon;v}|P) = \varepsilon^{2\gamma}\int Q^{\varepsilon;v}(d\omega)\log E^{\varepsilon;v}(T,\omega)\\
&= \varepsilon^{2\gamma}\int Q^{\varepsilon;v}(d\omega)\big(M^{\varepsilon;v}(T,\omega) - [M^{\varepsilon;v}, M^{\varepsilon;v}](T,\omega)\big)\\
&\quad + \frac{\varepsilon^{2\gamma}}{2}\int Q^{\varepsilon;v}(d\omega)[M^{\varepsilon;v}, M^{\varepsilon;v}](T,\omega) \le I(\mu) \qquad (3.22)
\end{aligned}$$

where in the last line we used Girsanov theorem, stating that $M^{\varepsilon;v} - [M^{\varepsilon;v}, M^{\varepsilon;v}]$ is a $Q^{\varepsilon,v}$-martingale and it has therefore vanishing expectation, and (3.21).

By (3.22), Lemma 3.6 and entropy inequality, the sequence $\{\mathbf{Q}^{\varepsilon;v}\}$ is tight in $\mathcal{P}(\mathcal{M})$, and in view of (3.22) it remains to show that any limit point of $\{\mathbf{Q}^{\varepsilon;v}\}$ is concentrated on $\{\delta_v\}$. Let $\mathbb{Q}^{\varepsilon;v} := Q^{\varepsilon;v} \circ (u^\varepsilon)^{-1} \in \mathcal{P}\big(C\big([0,T];U\big)\big)$; we will show

$$\lim_{\varepsilon} \mathbb{E}^{\mathbb{Q}^\varepsilon}\big(\sup_t \|u(t) - v(t)\|_{L_1(\mathbb{T})}\big) = 0 \qquad (3.23)$$

which is easily seen to imply the required convergence of $\{\mathbf{Q}^\varepsilon\}$. Since $\mathbb{Q}^{\varepsilon;v}$ is absolutely continuous with respect to $\mathbb{P}^\varepsilon$, it is concentrated on $C\big([0,T];U\big) \cap L_2\big([0,T];H^1(\mathbb{T})\big)$ and by Girsanov theorem it is a solution to the martingale problem associated with the stochastic partial differential equation in the unknown u

$$\begin{aligned}
du &= \Big[-\nabla \cdot f(u) + \frac{\varepsilon}{2}\nabla \cdot \big[D(u)\nabla u - a(u)\big((\jmath^\varepsilon * \jmath^\varepsilon) * (a(v)\nabla \Psi^v)\big)\big]\Big]\, dt\\
&\quad + \varepsilon^\gamma\, \nabla \cdot \big[a(u)(\jmath^\varepsilon * dW)\big]\\
u(0,x) &= u_0^\varepsilon(x) \qquad (3.24)
\end{aligned}$$

where we used the same notation of (2.2). Note that Ψ^v is twice continuously differentiable, since $a(v)^2$ is strictly positive and (3.19) can be regarded as an elliptical equation for Ψ^v with smooth data. Therefore by Lemma 3.3 applied with $E^\varepsilon = \jmath^\varepsilon * \jmath^\varepsilon * [a(v)\nabla\Psi^v]$ we have that $\mathbb{E}^{\mathbb{Q}^{\varepsilon;v}}\big(\varepsilon\|\nabla u\|^2_{L_2([0,T]\times\mathbb{T})}\big)$ is bounded uniformly in ε. By (3.19) and (3.24), we can then apply Lemma 3.4 with: $v^\varepsilon(u)(t,x) = v(t,x)$, $G_1^\varepsilon(u,v)(t,x) = \frac{\varepsilon}{2}\nabla \cdot \big[D(v)\nabla v\big]$, $G_2^\varepsilon(u,v) = 0$ and

$G_3^\varepsilon(u,v)(t,x) = [a(v)-a(u)]\big[a(v)\Psi^v - \jmath^\varepsilon * \jmath^\varepsilon * [a(v)\Psi^v]\big]$. Since v and Ψ^v are smooth, the hypotheses of Lemma 3.4 hold and we thus obtain (3.23). $\square$

Proof of Corollary 2.5. The corollary is an immediate consequence of the contraction principle [8, Theorem 4.4.1] applied to the continuous map $\mathcal{M} \ni \mu \mapsto \mu(\imath) \in C\big([0,T];U\big)$. If $\mu \in \mathcal{M}$ is such that $\mathcal{I}(\mu) < \infty$, then there exists $\Phi \in L_2\big([0,T]\times\mathbb{T}\big)$ such that $\partial_t \mu(\imath) = -\nabla\Phi$ holds weakly, and thus we have for $u \in C\big([0,T];U\big)$ and any Φ in the above class

$$\begin{aligned}\inf_{\mu\in\mathcal{M},\,\mu(\imath)=u} I(\mu) = &\inf_{\mu\in\mathcal{M},\,\mu(\imath)=u}\\ &\sup_{\varphi\in C^\infty([0,T]\times\mathbb{T})}\Big\{\langle\langle \Phi-\mu(f),\nabla\varphi\rangle\rangle - \frac{1}{2}\langle\langle \mu(a^2)\nabla\varphi,\nabla\varphi\rangle\rangle\Big\}\\ = &\int dt\,\inf_{c\in\mathbb{R}}\int dx \inf_{\mu\in\mathcal{M},\,\mu(\imath)=u}\frac{\big[\mu_{t,x}(f)-\Phi(t,x)-c\big]^2}{\mu_{t,x}(a^2)}\end{aligned}$$

Since the function Φ satisfying $\partial_t\mu(\imath) = -\nabla\Phi$ are defined up to a measurable additive function of t, the optimization over c can be replaced by an optimization over Φ, namely

$$\begin{aligned}\inf_{\mu\in\mathcal{M},\,\mu(\imath)=u} I(\mu) = &\inf_{\Phi\in L_2([0,T]\times\mathbb{T}),\nabla\Phi=-\partial_t u}\\ &\inf_{\mu\in\mathcal{M},\,\mu(\imath)=u}\int dt\,dx\,\frac{\big[\mu_{t,x}(f)-\Phi(t,x)\big]^2}{\mu_{t,x}(a^2)}\end{aligned}$$

which coincides with $I(u)$. $\square$

3.3. **Large deviations with speed $\varepsilon^{-2\gamma+1}$.** The next statement follows easily from entropy inequality (see also the introduction of [18] for further details).

Lemma 3.8. *Let $\mathcal{X}$ be a Polish space and $\{\mathbb{P}^\varepsilon\} \subset \mathcal{P}(\mathcal{X})$. The following are equivalent:*

- *(i) $\{\mathbb{P}^\varepsilon\}$ is exponentially tight with speed $\varepsilon^{-2\gamma+1}$.*
- *(ii) If a sequence $\{\mathbb{Q}^\varepsilon\} \subset \mathcal{P}(\mathcal{X})$ is such that $\overline{\lim}_\varepsilon\, \varepsilon^{2\gamma-1}\mathrm{Ent}(\mathbb{Q}^\varepsilon|\mathbb{P}^\varepsilon) < +\infty$, then $\{\mathbb{Q}^\varepsilon\}$ is tight.*

Let $\mathbb{Q} \in \mathcal{P}(\mathcal{X})$. For $\Psi : [0,T]\times\mathcal{X} \to C^\infty([0,T]\times\mathbb{T})$ a predictable process, let

$$\|\Psi\|^2_{\mathcal{D}^\varepsilon(\mathbb{Q})} := \int \mathbb{Q}(du)\big\|\jmath^\varepsilon * [a(u)\nabla\Psi(u)]\big\|^2_{L_2([0,T]\times\mathbb{T})} \in [0,+\infty]$$

We let $\mathcal{D}^\varepsilon(\mathbb{Q})$ be the Hilbert space obtained by identifying and completing the set of predictable processes $\Psi : [0,T]\times\mathcal{X} \to C^\infty([0,T]\times\mathbb{T})$ such that $\|\cdot\|_{\mathcal{D}^\varepsilon(\mathbb{Q})} < +\infty$ with respect to this seminorm.

Lemma 3.9. *Let* $\varepsilon > 0$ *and* $\mathbb{Q} \in \mathcal{P}(\mathcal{X})$ *be such that* $\mathrm{Ent}(\mathbb{Q}|\mathbb{P}^\varepsilon) < +\infty$*. Then there exists* $\Psi \in \mathcal{D}^\varepsilon(\mathbb{Q})$ *such that* $\mathbb{Q}$ *is a martingale solution to the Cauchy problem in the unknown* u

$$\begin{aligned} du &= \big(-\nabla \cdot f(u) + \frac{\varepsilon}{2}\nabla \cdot \big(D(u)\nabla u\big) - \varepsilon^\gamma \nabla \cdot \big[a(u)\jmath^\varepsilon * \jmath^\varepsilon * [a(u)\nabla\Psi(u)]\big]\big)dt \\ &\qquad + \varepsilon^\gamma\, \nabla \cdot \big[a(u)(\jmath^\varepsilon * dW)\big] \\ u(0,x) &= u_0^\varepsilon(x) \end{aligned} \tag{3.25}$$

and $\mathrm{Ent}(\mathbb{Q}|\mathbb{P}^\varepsilon) \geq \frac{1}{2}\|\Psi\|^2_{\mathcal{D}^\varepsilon(\mathbb{Q})}$.

Proof. Since $\mathbb{Q}$ is absolutely continuous with respect to $\mathbb{P}^\varepsilon$, there exists a continuous local $\mathbb{P}^\varepsilon$-martingale N on $\mathcal{X}$ such that

$$\mathbb{Q}(du) = \exp\big(N(T,u) - \frac{1}{2}\big[N,N\big](T,u)\big)\mathbb{P}^\varepsilon(du)$$

and $\mathrm{Ent}\big(\mathbb{Q}|\mathbb{P}^\varepsilon\big) = \mathbb{E}^{\mathbb{Q}}\big(N(T) - \frac{1}{2}\big[N,N\big](T)\big) = \frac{1}{2}\mathbb{E}^{\mathbb{Q}}\big(\big[N,N\big](T)\big)$ by Girsanov theorem.

It is easy to see that, as φ runs in $C^\infty([0,T]\times\mathbb{T})$, the family of maps (defined $\mathbb{P}$ a.s.)

$$\begin{aligned} [0,T] \times \chi \ni (t,u) \mapsto \langle M(t,u),\varphi\rangle &:= \langle u(t),\varphi(t)\rangle - \langle u(0),\varphi(0)\rangle \\ &- \int_0^t ds\, \langle u, \partial_t\varphi\rangle - \langle f(u) - \frac{1}{2}D(u)\nabla u, \nabla\varphi\rangle \in \mathbb{R} \end{aligned}$$

generates the standard filtration of $\mathcal{X}$. Therefore the martingale N is adapted to $\{\langle M,\varphi\rangle\}$, and reasoning as in [19, Lemma 4.2], there exists a predictable process Ψ on $\mathcal{X}$ and a martingale $\tilde{N}$ such that

$$N(t) = \int_0^t \langle \Psi, dM\rangle + \tilde{N}(t)$$

and

$$\big[\tilde{N}, \langle M,\varphi\rangle\big](T,u) = 0 \qquad \text{for all } \varphi \in C^\infty(\mathbb{T}), \text{ for } \mathbb{P} \text{ a.e. } u. \tag{3.26}$$

In particular

$$\begin{aligned} \mathbb{E}^{\mathbb{Q}}\big(\big[N,N\big](T)\big) &= \mathbb{E}^{\mathbb{Q}}\big(\big[\int_0^\cdot \langle\Psi,dM\rangle, \int_0^\cdot \langle\Psi,dM\rangle\big](T)\big) + \mathbb{E}^{\mathbb{Q}}\big(\big[\tilde{N},\tilde{N}\big](T)\big) \\ &\geq \mathbb{E}^{\mathbb{Q}}\big(\big\|\jmath^\varepsilon * [a(u)\nabla\Psi(u)]\big\|^2_{L_2([0,T]\times\mathbb{T})}\big) \end{aligned}$$

Therefore $\mathrm{Ent}(\mathbb{Q}|\mathbb{P}^\varepsilon) \geq \frac{1}{2}\|\Psi\|_{\mathcal{D}^\varepsilon(\mathbb{Q})}$ and (3.25) follows by Girsanov theorem and (3.26). It is immediate to see that both the bound on the relative entropy $\mathrm{Ent}(\mathbb{Q}|\mathbb{P}^\varepsilon)$ and the Girsanov term in (3.25) are compatible with the identification induced by the seminorm $\|\cdot\|_{\mathcal{D}^\varepsilon(\mathbb{Q})}$, and thus one can identify Ψ with an element in $\mathcal{D}^\varepsilon(\mathbb{Q})$. $\square$

Lemma 3.10. *Under the same hypotheses of Theorem 2.8 item (i), there exists a sequence $\{K_\ell\}$ of compact subsets of $\mathcal{X}$ such that*

$$\lim_{\ell} \overline{\lim_{\varepsilon}}\, \varepsilon^{2\gamma-1} \log \mathbb{P}^\varepsilon(K_\ell) = -\infty$$

Proof. In view of Lemma 3.8, we will prove that if $\mathbb{Q}^\varepsilon \subset \mathcal{P}(\mathcal{X})$ is a sequence with $\varepsilon^{2\gamma-1}\mathrm{Ent}(\mathbb{Q}^\varepsilon|\mathbb{P}^\varepsilon) \le C$ for some $C \ge 0$ independent of ε, then $\mathbb{Q}^\varepsilon$ is tight. By Lemma 3.9, there exists a sequence $\Psi^\varepsilon \in \mathcal{D}^\varepsilon(\mathbb{Q}^\varepsilon)$ such that

$$\frac{\varepsilon^{-1}}{2}\|\Psi^\varepsilon\|^2_{\mathcal{D}^\varepsilon(\mathbb{Q}^\varepsilon)} \le \varepsilon^{2\gamma-1}\mathrm{Ent}(\mathbb{Q}^\varepsilon|\mathbb{P}^\varepsilon) \le C \tag{3.27}$$

and $\mathbb{Q}^\varepsilon$ is a martingale solution to the Cauchy problem in the unknown u

$$\begin{aligned} du = \Big(&- \nabla \cdot f(u) + \frac{\varepsilon}{2}\nabla \cdot \big(D(u)\nabla u\big) \\ &- \nabla \cdot \big[a(u)\jmath^\varepsilon * \jmath^\varepsilon * [a(u)\nabla\Psi^\varepsilon(u)]\big]\Big)\, dt + \varepsilon^\gamma\, \nabla \cdot \big[a(u)(\jmath^\varepsilon * dW)\big] \\ u(0,x) = u_0^\varepsilon(x) & \end{aligned} \tag{3.28}$$

For $\varepsilon > 0$, we next define ($\mathbb{P}^\varepsilon$ a.s.) the predictable map $v^\varepsilon : \mathcal{X} \to \mathcal{X}$ as the solution to the parabolic Cauchy problem

$$\begin{aligned} &\partial_t v = -\nabla \cdot f(v) + \frac{\varepsilon}{2}\nabla \cdot \big(D(v)\nabla v\big) - \nabla \cdot \big[a(v)\jmath^\varepsilon * \jmath^\varepsilon * [a(u)\nabla\Psi^\varepsilon(u)]\big] \\ &v(0,x) = u_0(x) \end{aligned} \tag{3.29}$$

It is easily seen that, for $\mathbb{P}^\varepsilon$ a.e. u, (3.29) admits a unique solution $v^\varepsilon(u) \in \mathcal{X} \cap L_2\big([0,T]; H^1(\mathbb{T})\big)$, and that the definition of v^ε is compatible with the equivalence relation for Ψ^ε in the definition of $\mathcal{D}^\varepsilon(\mathbb{Q}^\varepsilon)$. By (3.29) and Young inequality for convolutions we also have

$$\begin{aligned} I_\varepsilon(v^\varepsilon(u)) &= \frac{1}{2}\big\|\jmath^\varepsilon * \jmath^\varepsilon * [a(u)\nabla\Psi^\varepsilon(u)]\big\|^2_{L_2([0,T]\times\mathbb{T})} \\ &\le \frac{1}{2}\big\|\jmath^\varepsilon * [a(u)\nabla\Psi^\varepsilon(u)]\big\|^2_{L_2([0,T]\times\mathbb{T})} \end{aligned} \tag{3.30}$$

where $I_\varepsilon : \mathcal{X} \cap L_2\big([0,T]; H^1(\mathbb{T})\big) \to [0,+\infty]$ is defined as

$$\begin{aligned} I_\varepsilon(v) \;:=\; \sup_{\varphi \in C^\infty([0,T]\times\mathbb{R})} \Big[&\langle v(T), \varphi(T)\rangle - \langle u_0, \varphi(0)\rangle \\ &- \langle\langle v, \partial_t\varphi\rangle\rangle + \langle\langle f(v) - \frac{1}{2}D(v)\nabla v, \nabla\varphi\rangle\rangle - \frac{1}{2}\langle\langle a(v)^2\nabla\varphi, \nabla\varphi\rangle\rangle\Big] \end{aligned}$$

Therefore taking the $\mathbb{E}^{\mathbb{Q}^\varepsilon}$ expectation in (3.30), multiplying by ε^{-1} and using (3.27)

$$E^{\mathbb{Q}^\varepsilon}\big(\varepsilon^{-1} I_\varepsilon(v^\varepsilon(u))\big) \le \frac{\varepsilon^{-1}}{2}\|\Psi^\varepsilon\|^2_{\mathcal{D}^\varepsilon(\mathbb{Q}^\varepsilon)} \le C \tag{3.31}$$

Minor adaptations of the proof of [3, Theorem 2.5] imply that for each $\ell > 0$ there exist $\varepsilon_0(\ell) > 0$ and a compact $K_\ell \subset \mathcal{X}$ such that

$$\cup_{\varepsilon \le \varepsilon_0(\ell)} \big\{ v \in \mathcal{X} \cap L_2\big([0,T]; H^1(\mathbb{T})\big) \,:\, \varepsilon^{-1} I_\varepsilon(v) \le \ell \big\} \subset K_\ell \tag{3.32}$$

(3.31) and (3.32) imply that the the sequence $\{\mathbb{Q}^\varepsilon \circ (v^\varepsilon)^{-1}\} \subset \mathcal{P}(\mathcal{X})$ is tight in $\mathcal{X}$, since by Chebyshev inequality

$$\big(\mathbb{Q}^\varepsilon \circ (v^\varepsilon)^{-1}\big)(K_\ell^c) \le C \ell^{-1}$$

By Lemma 3.3 (applied to $\mathbb{P}^\varepsilon$ with $E^\varepsilon \equiv 0$) and entropy inequality, we have

$$\lim_{\ell \to +\infty} \overline{\lim_{\varepsilon}}\, \mathbb{Q}^\varepsilon \big(\varepsilon \|\nabla u\|^2_{L_2([0,T]\times\mathbb{T})} \ge \ell \big) = 0$$

Therefore, in view of (3.28) and (3.29) we can apply Lemma 3.4 to $\mathbb{Q}^\varepsilon$ with $G_1(u,v) = 0$, $G_2(u,v) = 0$, $G_3(u,v) = [a(v) - a(u)]\big[\jmath^\varepsilon * \jmath^\varepsilon * [a(u)\nabla \Psi^\varepsilon(u)]\big]$. Indeed, since (3.27) holds, the hypotheses of Lemma 3.4 are easily satisfied. We then gather for each $\delta > 0$

$$\lim_{\varepsilon} \mathbb{Q}^\varepsilon \big(\sup_t \|u - v^\varepsilon(u)\|_{L_1(\mathbb{T})} \ge \delta \big) = 0$$

which implies, together with the tightness of $\{\mathbb{Q}^\varepsilon \circ (v^\varepsilon)^{-1}\}$ proved above, the tightness of $\{\mathbb{Q}^\varepsilon\}$. $\square$

Proof of Theorem 2.8: upper bound. Let $\mathcal{W} \subset \mathcal{X}$ be the set of weak solutions to (2.3). Let $K \subset \mathcal{X}$ be compact, and set $\mathcal{K} := \{\mu \in \mathcal{M} \,:\, \mu = \delta_u, \text{ for some } u \in K\}$. $\mathcal{K}$ is compact in $\mathcal{M}$, since $\mathcal{X}$ is equipped with the topology induced by the map $\mathcal{X} \ni u \mapsto \delta_u \in \mathcal{M}$. If $K \cap \mathcal{W} = \emptyset$, then $\inf_{\mu \in \mathcal{K}} \mathcal{I}(\mu) > 0$ as $\mathcal{I}$ vanishes only on measure-valued solutions to (2.3). In particular by Theorem 2.4 item (i)

$$\overline{\lim_{\varepsilon}}\, \varepsilon^{2\gamma-1} \log \mathbb{P}^\varepsilon(K) = \overline{\lim_{\varepsilon}}\, \varepsilon^{2\gamma-1} \log \mathbf{P}^\varepsilon(\mathcal{K}) = -\infty$$

Then, since $\mathcal{W}$ is closed in $\mathcal{X}$ and Lemma 3.10 holds, we need to prove the large deviations upper bound for $\{\mathbb{P}^\varepsilon\}$ only for compact sets $K \subset \mathcal{W} \subset \mathcal{X}$.

Let (ϑ, Q) be an entropy sampler–entropy sampler flux pair. Recall the definition of the martingale $N^{\varepsilon;\vartheta}$ in Lemma 3.1, and consider its stochastic

exponential

$$
\begin{aligned}
E^{\varepsilon;\vartheta}(t,u) &:= \exp\big(N^{\varepsilon,\vartheta}(t,u) - \frac{1}{2}\big[N^{\varepsilon,\vartheta}, N^{\varepsilon,\vartheta}\big](t,u)\big) \\
&= \exp\Big\{ \int dx\, \vartheta(u(t),t,x) - \int dx\, \vartheta(u_0,0,x) \\
&\quad - \int_0^t ds \int dx\, \big[\big(\partial_s \vartheta\big)\big(u(s,x),s,x\big) + \big(\partial_x Q\big)\big(u(s,x),s,x\big)\big] \\
&\quad + \int_0^t ds\, \Big[\frac{\varepsilon}{2}\langle \vartheta''(u)\nabla u, D(u)\nabla u\rangle + \frac{\varepsilon}{2}\langle \partial_x \vartheta'(u), D(u)\nabla u\rangle \\
&\qquad\qquad - \frac{\varepsilon^{2\gamma}}{2}\|\nabla \jmath^\varepsilon\|^2_{L_2(\mathbb{T})}\langle \vartheta''(u)a(u), a(u)\rangle \\
&\qquad\qquad - \frac{\varepsilon^{2\gamma}}{2}\|\jmath^\varepsilon\|^2_{L_2(\mathbb{T})}\langle \vartheta''(u)\nabla u, [a'(u)]^2\nabla u\rangle\Big] \\
&\quad - \frac{\varepsilon^{2\gamma}}{2}\int_0^t ds\, \big\langle a(u)^2\big[\vartheta''(u)\nabla u + \partial_x\vartheta'(u)\big], \vartheta''(u)\nabla u + \partial_x \vartheta'(u)\big\rangle\Big\}
\end{aligned}
$$

$E^{\varepsilon;\vartheta}$ is a continuous strictly positive $\mathbb{P}^\varepsilon$-supermartingale starting at 1. For $\ell > 0$ let

$$
B^\ell := \{u \in \mathcal{X} \cap L_2\big([0,T]; H^1(\mathbb{T})\big) \,:\, \|\nabla u\|^2_{L_2([0,T]\times\mathbb{T})} \le \ell\}
$$

Recall that $\mathcal{W}$ is the set of weak solutions to (1.3). Given a Borel subset $A \subset \mathcal{W}$ we have, for C, ε_0 as in Lemma 3.3 (applied with $E^\varepsilon \equiv 0$) and $\ell > C$, $\varepsilon \le \varepsilon_0$

$$
\begin{aligned}
\mathbb{P}^\varepsilon(A) &\le \mathbb{E}^{P^\varepsilon}\big(E^{\varepsilon;\frac{\vartheta}{\varepsilon^{2\gamma-1}}}(T,u)[E^{\varepsilon;\frac{\vartheta}{\varepsilon^{2\gamma-1}}}(T,u)]^{-1}\mathbb{1}_{A\cap B^{\ell/\varepsilon}}(u)\big) + \mathbb{P}^\varepsilon(B^{\ell/\varepsilon}) \\
&\le \sup_{u \in A\cap B^{\ell/\varepsilon}} [E^{\varepsilon;\frac{\vartheta}{\varepsilon^{2\gamma-1}}}(T,v)]^{-1} + \exp\big(-\frac{(\ell-C)^2}{C\varepsilon^{2\gamma-1}(\ell+1)}\big) \qquad (3.33)
\end{aligned}
$$

where in the last line we used the supermartingale property of $E^{\varepsilon;\vartheta}$ and Lemma 3.3. Since

$$
\begin{aligned}
\varepsilon^{2\gamma-1}\log E^{\varepsilon;\frac{\vartheta}{\varepsilon^{2\gamma-1}}}(T,u) = &-\int dx\, \vartheta(u_0(x),0,x) \\
&-\int ds\, dx\, \big[\big(\partial_s\vartheta\big)\big(u(s,x),s,x\big) + \big(\partial_x Q\big)\big(u(s,x),s,x\big)\big] \\
&+\frac{\varepsilon}{2}\langle\langle \vartheta''(u)\nabla u, D(u)\nabla u\rangle\rangle + \frac{\varepsilon}{2}\langle\langle \partial_x\vartheta'(u), D(u)\nabla u\rangle\rangle \\
&-\frac{\varepsilon^{2\gamma}}{2}\|\nabla\jmath^\varepsilon\|^2_{L_2(\mathbb{T})}\langle\langle \vartheta''(u)a(u), a(u)\rangle\rangle \\
&-\frac{\varepsilon^{2\gamma}}{2}\|\jmath^\varepsilon\|^2_{L_2(\mathbb{T})}\langle\langle \vartheta''(u)\nabla u, [a'(u)]^2\nabla u\rangle\rangle \\
&-\frac{\varepsilon}{2}\langle\langle a(u)^2\vartheta''(u)\nabla u, \vartheta''(u)\nabla u\rangle\rangle - \frac{\varepsilon}{2}\langle\langle a(u)^2\partial_x\vartheta'(u), \partial_x\vartheta'(u)\rangle\rangle \\
&-\varepsilon\langle\langle a(u)^2\vartheta''(u)\nabla u, \partial_x\vartheta'(u)\rangle\rangle
\end{aligned}
$$

by Cauchy-Schwartz inequality, for each $u \in B^{\ell/\varepsilon}$

$$
\begin{aligned}
\varepsilon^{2\gamma-1}\log E^{\varepsilon;\frac{\vartheta}{\varepsilon^{2\gamma-1}}}(T,u) \geq &-\int dx\, \vartheta(u_0(x),0,x) \\
&-\int ds\, dx\, \big[\big(\partial_s\vartheta\big)\big(u(s,x),s,x\big) + \big(\partial_x Q\big)\big(u(s,x),s,x\big)\big] \\
&+\frac{\varepsilon}{2}\langle\langle \vartheta''(u)\nabla u, \big(D(u)-a(u)^2\vartheta''(u)\big)\nabla u\rangle\rangle - C_\vartheta\sqrt{\varepsilon\ell} \\
&-C_\vartheta\varepsilon^{2\gamma}\|\nabla\jmath^\varepsilon\|^2_{L_2(\mathbb{T})} - C_\vartheta\varepsilon^{2\gamma-1}\ell\|\jmath^\varepsilon\|^2_{L_2(\mathbb{T})} - C_\vartheta\varepsilon - \sqrt{\varepsilon\ell}C_\vartheta \quad (3.34)
\end{aligned}
$$

for a suitable constant $C_\vartheta > 0$ depending only on ϑ, D and a. The key point now is that, if the entropy sampler ϑ satisfies

$$
a(u)^2\vartheta''(u,t,x) \leq D(u) \qquad \forall\, u\in[0,1],\ t\in[0,T],\ x\in\mathbb{T} \tag{3.35}
$$

then the term $\langle\langle \vartheta''(u)\nabla u, \big(D(u)-a(u)^2\vartheta''(u)\big)\nabla u\rangle\rangle$ in (3.34) is positive. Namely, the largest term in the quadratic variation of $N^{\varepsilon;\vartheta}$ is controlled by the positive parabolic term associated with the deterministic diffusion. Therefore taking the limit $\varepsilon \to 0$ in (3.34), by the hypotheses assumed on $\jmath^\varepsilon$, for each entropy sampler ϑ satisfying (3.35) and each $u \in B^{\ell/\varepsilon}$

$$
\begin{aligned}
\overline{\lim_{\varepsilon}}\, \varepsilon^{2\gamma-1}\log E^{\varepsilon;\frac{\vartheta}{\varepsilon^{2\gamma-1}}}(T,u) \geq &-\int dx\, \vartheta(u_0(x),0,x) \\
&-\int ds\, dx\, \big[\big(\partial_s\vartheta\big)\big(u(s,x),s,x\big) + \big(\partial_x Q\big)\big(u(s,x),s,x\big)\big] \quad (3.36)
\end{aligned}
$$

We now take the logarithm of (3.33) and multiply it by $\varepsilon^{2\gamma-1}$. Taking the limits $\varepsilon \to 0$, then $\ell \to +\infty$, and using (3.36), we have for each ϑ satisfying

(3.35)

$$
\begin{aligned}
\overline{\lim_{\varepsilon}}\, \varepsilon^{2\gamma-1} \log \mathbb{P}^\varepsilon(A) \le -\inf_{u\in A} \Big\{ -\int dx\, \vartheta(u_0(x),0,x) \\
-\int ds\, dx\, \big[\big(\partial_s \vartheta\big)\big(u(s,x),s,x\big) + \big(\partial_x Q\big)\big(u(s,x),s,x\big)\big]\Big\} \le -\inf_{u\in A} \sup_{\vartheta} P_{\vartheta,u}
\end{aligned}
$$

where we have applied the definition (2.8) of $P_{\vartheta,u}$. Note that the map $\mathcal{X} \ni u \mapsto P_{\vartheta,u} \in \mathbb{R}$ is lower semicontinuous. Applying the minimax lemma, we gather for a compact set $K \subset \mathcal{W}$

$$
\overline{\lim_{\varepsilon}}\, \varepsilon^{2\gamma-1} \mathbb{P}^\varepsilon(K) \le -\inf_{u\in K} \sup_{\vartheta} P_{\vartheta,u}
$$

where the supremum is taken over the entropy samplers ϑ satisfying (3.35). It is easy to see that a weak solution u to (2.3) such that $\sup_\vartheta P_{\vartheta,u} < +\infty$ is indeed an entropy-measure solution $u \in \mathcal{E}$, and $\sup_\vartheta P_{\vartheta,u} = H(u)$. $\square$

Proof of Theorem 2.8: lower bound. We will use the entropy method suggested by Lemma 3.7, as we did in the proof of Theorem 2.4 item (ii). Recall the Definition 2.7 of $\mathcal{S}$. Given $v \in \mathcal{S}$, we need to show that there exists a sequence $\{\mathbb{Q}^{\varepsilon;v}\} \subset \mathcal{P}(\mathcal{X})$ such that $\overline{\lim}\, \varepsilon^{2\gamma-1} \mathrm{Ent}(\mathbb{Q}^{\varepsilon;v}|\mathbb{P}^\varepsilon) \le H(u)$ and $\mathbb{Q}^\varepsilon \to \delta_v$ in $\mathcal{P}(\mathcal{X})$. The lower bound with rate $\overline{H}$ then follows by a standard diagonal argument.

With minor adaptations from Theorem 2.5 in [3], we have that the following statement holds.

Lemma 3.11. *For each sequence $\beta_\varepsilon \to 0$ and each $v \in \mathcal{S}$, there exist a sequence $\{w^\varepsilon\} \subset \mathcal{X} \cap L_2\big([0,T]; H^1(\mathbb{T})\big)$ and a sequence $\{\Psi^\varepsilon\} \subset L_2\big([0,T]; H^2(\mathbb{T})\big)$ such that:*

- (a) *$w^\varepsilon \to v$ in $\mathcal{X}$, and $w^\varepsilon(0,x) = u_0(x)$.*
- (b) *$\varepsilon \|\nabla w^\varepsilon\|^2_{L_2([0,T]\times\mathbb{T})} \le C$ for some $C > 0$ independent of ε.*
- (c) *$\overline{\lim}_\varepsilon \frac{\varepsilon^{-1}}{2} \langle\langle a(w^\varepsilon)^2 \nabla \Psi^\varepsilon, \nabla \Psi^\varepsilon \rangle\rangle = H(v)$.*
- (d) *$\beta_\varepsilon \big\|\nabla[a(w^\varepsilon)\nabla\Psi^\varepsilon]\big\|^2_{L_2([0,T]\times\mathbb{T})} \le C\, \varepsilon^{-1}$, for some $C > 0$ independent of ε.*
- (e) *The equation*

$$
\partial_t w^\varepsilon + \nabla \cdot f(w^\varepsilon) - \frac{\varepsilon}{2} \nabla \cdot \big(D(w^\varepsilon)\nabla w^\varepsilon\big) = -\nabla \cdot \big(a(w^\varepsilon)^2\, \nabla \Psi^\varepsilon\big)
$$

holds weakly.

We let $\beta_\varepsilon := \varepsilon^{-3/2} \|\jmath^\varepsilon - \mathbb{1}\|_{W^{-1,1}(\mathbb{T})}$, and let $\{w^\varepsilon\}$, $\{\Psi^\varepsilon\}$ be chosen correspondingly. Note that with this choice of β_ε and by the assumption on $\|\jmath^\varepsilon - \mathbb{1}\|_{W^{-1,1}(\mathbb{T})}$

$$
\lim_{\varepsilon} \varepsilon^{-2} \int_0^t ds\, \|\jmath^\varepsilon * \jmath^\varepsilon * [a(w^\varepsilon)\nabla\Psi^\varepsilon] - a(w^\varepsilon)\nabla\Psi^\varepsilon\|^2_{L_2(\mathbb{T})} = 0 \tag{3.37}
$$

We define the martingale $M^{\varepsilon;v}$ on Ω as

$$M^{\varepsilon;v}(t) := \varepsilon^{-\gamma} \int_0^t \langle \jmath^\varepsilon * [a(w^\varepsilon)\nabla\Psi^\varepsilon], dW\rangle$$

Then by Young inequality for convolutions:

$$\frac{1}{2}\big[M^{\varepsilon;v}, M^{\varepsilon;v}\big](T) \le \frac{\varepsilon^{-2\gamma}}{2}\langle\langle a(w^\varepsilon)^2\nabla\Psi^\varepsilon, \nabla\Psi^\varepsilon\rangle\rangle \tag{3.38}$$

In particular the stochastic exponential of $N^{\varepsilon;v}$ is a martingale on Ω, and we can define the probability measure $Q^{\varepsilon;v} \in \mathcal{P}(\Omega)$ as

$$Q^{\varepsilon;v}(d\omega) := \exp\big(N^{\varepsilon;v}(T,\omega) - \frac{1}{2}\big[N^{\varepsilon;v}, N^{\varepsilon;v}\big](T,\omega)\big)P(d\omega)$$

and $\mathbb{Q}^{\varepsilon;v} := Q^{\varepsilon;v} \circ (u^\varepsilon)^{-1} \in \mathcal{P}(\mathcal{X})$, where $u^\varepsilon : \Omega \to \mathcal{X}$ is the solution to (2.2). Reasoning as in (3.22), and using (3.38) and property (c) in Lemma 3.11

$$\begin{aligned}
\overline{\lim_{\varepsilon}}\, \varepsilon^{2\gamma-1}\mathrm{Ent}(\mathbb{Q}^{\varepsilon;v}|\mathbb{P}^{\varepsilon;v}) &\le \overline{\lim_{\varepsilon}}\, \varepsilon^{2\gamma-1}\mathrm{Ent}(Q^{\varepsilon;v}|P) \\
&= \overline{\lim_{\varepsilon}}\, \frac{\varepsilon^{2\gamma-1}}{2}\int Q^{\varepsilon;v}(d\omega)[M^{\varepsilon;v}, M^{\varepsilon;v}](T,\omega) \\
&\le \overline{\lim_{\varepsilon}}\, \frac{\varepsilon^{-1}}{2}\langle\langle a(w^\varepsilon)^2\nabla\Psi^\varepsilon, \nabla\Psi^\varepsilon\rangle\rangle = H(v) \qquad (3.39)
\end{aligned}$$

We next need to prove that $\mathbb{Q}^{\varepsilon;v}$ converges to δ_v in $\mathcal{P}(\mathcal{X})$ as $\varepsilon \to 0$. By Girsanov theorem $\mathbb{Q}^{\varepsilon;v}$ is a martingale solution to the stochastic Cauchy problem in the unknown u

$$\begin{aligned}
du &= \big[-\nabla\cdot f(u) + \frac{\varepsilon}{2}\nabla\cdot\big(D(u)\nabla u\big) - \nabla\cdot a(u)(\jmath * \jmath * (a(w^\varepsilon)\nabla\Psi^\varepsilon)\big]\,dt \\
&\quad + \varepsilon^\gamma\,\nabla\cdot\big[a(u)(\jmath^\varepsilon * dW)\big] \\
u(0,x) &= u_0^\varepsilon(x)
\end{aligned} \tag{3.40}$$

In view of property (a) in Lemma 3.11, it is enough to check that Lemma 3.4 holds with $v^\varepsilon(u)(t,x) = w^\varepsilon(t,x)$. Indeed, still by property (a) in Lemma 3.11 and the assumptions of this theorem, conditions (i) and (ii) in Lemma 3.4 are immediate. By property (e) in Lemma 3.11 and (3.40), $\mathbb{Q}^{\varepsilon;v}$ is a martingale solution to (3.12) with $G_1^\varepsilon \equiv 0$,

$$G_2^\varepsilon(u,w) = a(w)\big[\jmath^\varepsilon * \jmath^\varepsilon * [a(w^\varepsilon)\nabla\Psi^\varepsilon] - a(w^\varepsilon)\nabla\Psi^\varepsilon\big]$$

$$G_3^\varepsilon(u,w) = [a(w) - a(u)][\jmath * \jmath * (a(w^\varepsilon)\nabla\Psi^\varepsilon]$$

Therefore, in view of (3.37), condition (iii) in Lemma 3.4 is easily seen to hold. Condition (iv) is also immediate from the definition of G_3 and the bound on $\mathbb{Q}^{\varepsilon;v}\big(\varepsilon\|\nabla u\|_{L_2([0,T]\times\mathbb{T})} > \ell\big)$ provided by the application of Lemma 3.3 for $\mathbb{P}^\varepsilon$ (thus with $E^\varepsilon \equiv 0$), the entropy bound (3.39), and the usual entropy inequality. $\square$

Appendix A. Existence and uniqueness results for fully nonlinear parabolic SPDEs with conservative noise

In this appendix, we are concerned with existence and uniqueness results for the Cauchy problem in the unknown $u \equiv u(t,x)$, $t \in [0,T]$, $x \in \mathbb{T}$

$$\begin{aligned} du &= \big[-\nabla \cdot f(u) + \frac{1}{2} \nabla \cdot \big(D(u) \nabla u\big) \big]\, dt + \nabla \cdot \big[a(u) (\jmath * dW)\big] \\ u(0,x) &= u_0(x) \end{aligned} \tag{A.1}$$

Although we assume the space-variable x to run on a one-dimensional torus $\mathbb{T}$, it is not difficult to extend the results given below to the case $x \in \mathbb{T}^d$ or $x \in \mathbb{R}^d$ for $d \geq 1$.

Let W be an $L_2(\mathbb{T})$–valued cylindrical Brownian motion on a given standard filtered probability space $\big(\Omega, \mathfrak{F}, \{\mathfrak{F}_t\}_{0 \le t \le T}, P\big)$. Hereafter we set

$$Q(v) := a'(v)^2 \|\jmath\|^2_{L_2(\mathbb{T})}$$

We will assume the following hypotheses:

- **A1)** f and D are uniformly Lipschitz on $\mathbb{R}$.
- **A2)** $a \in C^2(\mathbb{R})$ is uniformly bounded.
- **A3)** $\jmath \in H^1(\mathbb{T})$ and, with no loss of generality, $\int dx\, |\jmath(x)| = 1$.
- **A4)** There exists $c > 0$ such that $D \geq Q + c$.
- **A5)** $u_0 : \Omega \to L_2(\mathbb{T})$ is $\mathfrak{F}_0$-Borel measurable and satisfies $\mathbb{E}^{\mathbb{P}}\big(\|u_0\|^2_{L_2(\mathbb{T})}\big) < +\infty$.

We introduce the Polish space $Y := C\big([0,T]; H^{-1}(\mathbb{T})\big) \cap L_2\big([0,T]; H^1(\mathbb{T})\big) \cap L_\infty\big([0,T]; L_2(\mathbb{T})\big)$. A probability measure $\bar{\mathbb{P}}$ on Y is a *martingale solution* to (A.1) iff the law of $u(0)$ under $\bar{\mathbb{P}}$ is the same of the law of u_0, and for each $\varphi \in C^\infty\big([0,T] \times \mathbb{T}\big)$

$$\begin{aligned} \langle M(t,u), \varphi \rangle := &\langle u(t), \varphi(t) \rangle - \langle u(0), \varphi(0) \rangle \\ &- \int_0^t ds\, \langle u, \partial_s \varphi \rangle + \langle f(u) - \frac{1}{2} D(u) \nabla u, \nabla \varphi \rangle \end{aligned} \tag{A.2}$$

is a continuous square-integrable martingale with respect to $\bar{\mathbb{P}}(du)$ with quadratic variation

$$\big[\langle M, \varphi \rangle, \langle M, \psi \rangle\big](t,u) = \int_0^t ds\, \langle \jmath * (a(u) \nabla \varphi), \jmath * (a(u) \nabla \psi) \rangle \tag{A.3}$$

We say that a progressively measurable process $u : \Omega \to Y$ is a *strong solution* to (A.1) iff $u(0) = u_0$ P-a.s. and for each $\varphi \in C^\infty\big([0,T] \times \mathbb{T}\big)$

$$\langle M, \varphi \rangle = - \int_0^t \langle \jmath * \big(a(u) \nabla \varphi\big), dW \rangle \tag{A.4}$$

In this appendix we prove

Theorem A.1. *Assume **A1)**–**A5)**. Then there exists a unique strong solution u to* (A.1) *in Y. Such a solution u admits a version in $C\big([0,T];L_2(\mathbb{T})\big)$. Furthermore, if u_0 takes values in $[0,1]$ and a is supported by $[0,1]$, then u takes values in $[0,1]$ a.s..*

By compactness estimates we will prove that there exists a solution to the martingale problem related to (A.1). Then we will provide pointwise uniqueness for (A.1) using a stability result similar to the one used in the proof of Lemma 3.4. By Yamada-Watanabe theorem we get the existence and uniqueness stated in Theorem A.1. We remark that assumption **A4)** is a key hypotheses in the proof of Theorem A.1, as it implies that the noise term is smaller than the second order parabolic term, thus allowing some a priori bounds. In general, one may expect nonexistence of the solution to (A.1) if such a condition fails, see [7, Example 7.21].

Lemma A.2. *Let $0 \le t' < t'' \le T$, let $u', v : \Omega \to L_2(\mathbb{T})$ be $\mathfrak{F}_{t'}$-measurable maps such that $\mathbb{E}^P\big(\||u'|+|v|+|\nabla v|\|^2_{L_2(\mathbb{T})}\big) < +\infty$. Then the stochastic Cauchy problem in the unknown w*

$$\begin{aligned} & dw = \big[-\nabla\cdot f(w) + \frac{1}{2}\nabla\cdot\big(D(v)\nabla w\big)\big]\,dt + \nabla\cdot\big[a(v)(\jmath * dW)\big] \\ & w(t',x) = u'(x) \end{aligned} \tag{A.5}$$

admits a unique strong solution u in $L_2\big([t',t''];H^1(\mathbb{T})\big)\cap C\big([t',t''],H^{-1}(\mathbb{T})\big)$ with probability 1. For each $t\in[t',t'']$, such a solution u satisfies

$$\begin{aligned} \langle u(t),u(t)\rangle + \int_{t'}^t ds \langle D(v)\nabla u,\nabla u\rangle = N(t,t') + \langle u',u'\rangle \\ + \int_{t'}^t ds\,\big[\langle Q(v)\nabla v,\nabla v\rangle + \|\nabla \jmath\|^2_{L_2(\mathbb{T})}\int dx\,a(v)^2\big] \end{aligned} \tag{A.6}$$

where $N(t,t') := -2\int_{t'}^t \langle \jmath * \big(a(v)\nabla u\big), dW\rangle$. Furthermore

$$\mathbb{E}^P\Big(\sup_{t\in[t',t'']}\|u(t)\|^2_{L_2(\mathbb{T})}\Big) < +\infty$$

Proof. Existence and uniqueness of the semilinear equation (A.5) are standard, see e.g. [7, Chap. 7.7.3]. Applying Itô formula to the function $L_2(\mathbb{T}) \ni w \mapsto \langle w,w\rangle \in \mathbb{R}$ we get (A.6). Note that by Burkholder-Davis-Gundy inequality [19, Theorem 4.4.1], Young and Cauchy-Schwarz inequalities, for suitable constants

$C, C' > 0$

$$\begin{aligned}
\mathbb{E}^P\big(\sup_{t\in[t',t'']} |N(t,t')|\big) &\le C\,\mathbb{E}^P\big(\big[N(\cdot,t'),N(\cdot,t')\big](t'')^{1/2}\big) \\
&= 2\,C\,\mathbb{E}^P\Big(\|\jmath * \big(a(v)\nabla u\big)\|_{L_2([t',t'']\times\mathbb{T})}\Big) \\
&\le 2\,C\,\mathbb{E}^P\Big(\|\big(a(v)\nabla u\big)\|_{L_2([t',t'']\times\mathbb{T})}\Big) \\
&\le C'\Big[\mathbb{E}^P\Big(\int_{t'}^{t''} ds\,\langle D(v)\nabla u, \nabla u\rangle\Big)\Big]^{1/2}
\end{aligned}$$

so that the bound on $\mathbb{E}^P\big(\sup_{t\in[t',t'']}\|u(t)\|^2_{L_2(\mathbb{T})}\big)$ is easily obtained by taking the supremum over t and the $\mathbb{E}^P$ expected values in (A.6). □

We next introduce a sequence $\{u^n\}$ of adapted processes in Y. We will gather existence of a weak solution to (A.1) by tightness of the laws $\{\mathbb{P}^n\}$ of such a sequence.

For $n \in \mathbb{N}$ and $i = 0,\dots,2^n$ let $t_i^n := i2^{-n}T$, and let $\{\imath^n\}$ be a sequence of smooth mollifiers on $\mathbb{T}$ such that $\lim_n 2^{-n}\|\imath^n\|^2_{L_1(\mathbb{T})} = 0$. We define a process u^n on Y and the auxiliary random functions $\{v_i^n\}_{i=0}^{2^n}$ on $\mathbb{T}$ as follows. For $i = 0$ we set

$$\begin{aligned}
u^n(0) &:= u_0 \\
v_0^n &:= \imath^n * u_0
\end{aligned}$$

and for $i = 1,\dots,2^n-1$ and $t \in [t_i^n, t_{i+1}^n]$, we let $u^n(t)$ be the solution to the problem (A.5) with $u' = u(t_i^n)$ and $v = v_i^n$, where for $i \ge 1$ we set

$$v_i^n := \frac{2^n}{T}\int_{t_{i-1}^n}^{t_i^n} ds\, u^n(s) \tag{A.7}$$

By Lemma A.2, these definitions are well-posed, and u^n is in Y with probability 1. We also define a sequence $\{v^n\}$ of cadlag processes in the Skorohod space $D\big([0,T);L_2(\mathbb{T})\big)$, by requiring

$$v^n(t) = v_i^n \text{for } t \in [t_i^n, t_{i+1}^n) \tag{A.8}$$

Lemma A.3. *There exists a constant $C > 0$ independent of n such that*

$$\mathbb{E}^P\Big(\sup_{t\in[0,T]}\|u^n(t)\|^2_{L_2(\mathbb{T})} + \|\nabla u^n\|^2_{L_2([0,T]\times\mathbb{T})}\Big) \le C \tag{A.9}$$

and for each $\varphi \in H^1(\mathbb{T})$ such that $\|\nabla\varphi\|^2_{L_2(\mathbb{T})} \le 1$, for each $\delta > 0$ and $r \in (0,1)$

$$P\big(\sup_{s,t\in[0,T]\,:|s-t|\le\delta}\big|\langle u^n(t) - u^n(s),\varphi\rangle\big| > r\big) \le C\,\delta\,r^{-2} \tag{A.10}$$

Furthermore for each $r > 0$

$$\lim_{n\to\infty} P\big(\|u^n - v^n\|_{L_2([0,T]\times\mathbb{T})} > r\big) = 0 \tag{A.11}$$

Proof. Writing Itô formula (A.6) for u^n in the intervals $[t_i^n, t_{i+1}^n]$ and summing over i, we get for each $t \in [0,T]$

$$\begin{aligned}\langle u^n(t), u^n(t)\rangle + \int_0^t ds\, \langle D(v^n)\nabla u^n, \nabla u^n\rangle = \langle u_0, u_0\rangle \\ + \int_0^t ds\, \big[\langle Q(v^n)\nabla v^n, \nabla v^n\rangle + \|\nabla \jmath\|^2_{L_2(\mathbb{T})} \int dx\, a(v^n)^2\big] + N^n(t)\end{aligned}$$

where, by the same means of Lemma A.2 and Doob's inequality, the martingale

$$N^n(t) := 2\int_0^t \langle \jmath * \big(a(v^n)\nabla u^n\big), dW\rangle$$

enjoys the bound

$$\mathbb{E}^P\big(\sup_{s\in[0,T]} |N^n(t)|^2\big) \le C_1\, \mathbb{E}^P\big(\|\nabla u^n\|^2_{L_2([0,T]\times\mathbb{T})}\big)$$

for some $C_1 > 0$ depending only on D and a. Note that, by the definition of v_i^n (A.7), hypotheses **A4)-A5)** and Young inequality for convolutions

$$\begin{aligned}&\int_0^t ds\, \langle Q(v^n)\nabla v^n, \nabla v^n\rangle \\ &\qquad \le C_2 \int_0^{t_1^n} ds\, \|\imath^n * u_0\|^2_{L_2(\mathbb{T})} + \int_0^t ds\, \langle Q(v^n)\nabla u^n, \nabla u^n\rangle \\ &\qquad \le 2^{-n} T\, C_2\, \|\imath^n\|^2_{L_1(\mathbb{T})} \|u_0\|^2_{L_2(\mathbb{T})} + \int_0^t ds\, \langle (D(v^n) - c)\nabla u^n, \nabla u^n\rangle\end{aligned}$$

for some constant C_2 depending only on a. Patching all together

$$\begin{aligned}&\mathbb{E}^P\big(\sup_{t\in[0,T]} \|u^n(t)\|^2_{L_2([0,T]\times\mathbb{T})} + c\langle\langle D(v^n)\nabla u^n, \nabla u^n\rangle\rangle\big) \\ &\quad \le \big(1 + 2^{-n} T\, C_2 \|\imath^n\|^2_{L_1(\mathbb{T})}\big)\, \mathbb{E}^P\big(\|u_0\|^2_{L_2(\mathbb{T})}\big) \\ &\qquad + C_1\, \mathbb{E}^P\big(\langle\langle D(v^n)\nabla u^n, \nabla u^n\rangle\rangle^{1/2}\big) + \|\nabla \jmath\|^2_{L_2(\mathbb{T})}\, \mathbb{E}^P\big(\|a(v^n)\|^2_{L_2([0,t]\times\mathbb{T})}\big)\end{aligned}$$

Since $2^{-n}\|\imath^n\|_{L_1(\mathbb{T})}$ was assumed bounded, and since the last term in the right hand side is bounded uniformly in n, it is not difficult to gather (A.9).

Since u satisfies (A.5) in each interval $[t_i^n, t_{i+1}^n]$

$$\begin{aligned}\big|\langle u^n(t) - u^n(s), \varphi\rangle\big| \;\le\; & C_3\big(1 + \|\nabla u^n\|_{L_2([0,T]\times\mathbb{T})}\big)\|\nabla\varphi\|_{L_2(\mathbb{T})}|t-s|^{1/2} \\ & + \big|\int_s^t \langle \jmath * (a(v)\nabla\varphi), dW\rangle\big|\end{aligned}$$

for a suitable constant C_3 depending only on f and D. (A.10) then follows from the first part of the lemma.

Since $v^n(t) = \imath^n * u_0$ for $t \in [0, t_1^n)$, the bound (A.9) implies

$$\lim_{n\to\infty} P\big(\|u^n - v^n\|_{L_2([0,t_1^n]\times\mathbb{T})} > r\big) = 0$$

for each $r > 0$. Therefore, still by (A.9), in order to prove (A.11), it is enough to show that for each $r,\, \ell > 0$

$$\lim_{n\to\infty} P\big(\|u^n - v^n\|_{L_2([t_1^n,T]\times\mathbb{T})} > r, \|\nabla u^n\|^2_{L_2([0,T]\times\mathbb{T})} \le \ell\big) = 0$$

Let $\kappa \in C^\infty(\mathbb{T})$ be such that $\int dx\, \kappa(x) = 1$, and that

$$\|\kappa - \mathrm{id}\|_{-1,1} := \sup\Big\{ \int dx\, \Big| \int dy\, \kappa(x-y)\varphi(y) - \varphi(x)\Big|, \\ \varphi \in C^\infty(\mathbb{T}) \,:\, \sup_x |\nabla\varphi(x)| \le 1\Big\} \le \frac{r}{2\ell} \quad \text{(A.12)}$$

It is immediate to see that such a κ exists. Then

$$\begin{aligned}\|u^n - v^n\|_{L_2([t_1^n,T]\times\mathbb{T})} &\le \|u^n - \kappa * u^n\|_{L_2([t_1^n,T]\times\mathbb{T})} \\ &\quad + \|v^n - \kappa * v^n\|_{L_2([t_1^n,T]\times\mathbb{T})} + \|\kappa * u^n - \kappa * v^n\|_{L_2([t_1^n,T]\times\mathbb{T})} \\ &\le \|\kappa - \mathrm{id}\|_{-1,1}\big[\|\nabla u^n\|_{L_2([t_1^n,T]\times\mathbb{T})} + \|\nabla v^n\|_{L_2([t_1^n,T]\times\mathbb{T})}\big] \\ &\quad + \|\kappa * (u^n - v^n)\|_{L_2([t_1^n,T]\times\mathbb{T})}\end{aligned}$$

where in the last inequality we used the Young inequality. By the definition (A.7)-(A.8) of v^n, $\|\nabla v^n\|^2_{L_2([t_1^n,T]\times\mathbb{T})} \le \|\nabla u^n\|^2_{L_2([0,T]\times\mathbb{T})}$. Moreover

$$\begin{aligned}&\int_{t_1^n}^T dt\, \|\kappa * (u^n - v^n)\|^2_{L_2(\mathbb{T})} \\ &\quad = \sum_{i=1}^{2^n-1} \int_{t_i^n}^{t_{i+1}^n} dt\, \Big\|\kappa * u^n(t) - \frac{2^n}{T}\int_{t_{i-1}^n}^{t_i^n} ds\, \kappa * u^n(s)\Big\|^2_{L_2(\mathbb{T})} \\ &\quad \le T \sup_{|t-s|\le 2^{-n+1}T} \|\kappa * (u^n(t) - u^n(s))\|^2_{L_2(\mathbb{T})}\end{aligned}$$

Therefore by (A.12)

$$\|u^n - v^n\|^2_{L_2([t_1^n,T]\times\mathbb{T})} \le \frac{r}{2\ell}\|\nabla u^n\|^2_{L_2([t_1^n,T]\times\mathbb{T})} \\ + T \sup_{|t-s|\le 2^{-n+1}T} \|\kappa * (u^n(t) - u^n(s))\|^2_{L_2(\mathbb{T})} \quad \text{(A.13)}$$

so that

$$\begin{aligned}&\lim_{n\to\infty} P\big(\|u^n - v^n\|_{L_2([t_1^n,T]\times\mathbb{T})} > r, \|\nabla u^n\|^2_{L_2([0,T]\times\mathbb{T})} \le \ell\big) \\ &\quad \le \varlimsup_{n\to\infty} P\big(\sqrt{T} \sup_{|t-s|\le 2^{-n+1}T} \|\kappa * (u^n(t) - u^n(s))\|_{L_2(\mathbb{T})} \ge r/2\big)\end{aligned}$$

which vanishes in view of (A.10). $\square$

We define $\mathbb{P}^n$ to be the law of u^n, namely $\mathbb{P}^n = P \circ (u^n)^{-1}$. In order to establish tightness of the sequence $\{\mathbb{P}^n\}$, the $\mathbb{P}^n$ will be regarded as probability measures on $C\big([0,T], H^{-1}(\mathbb{T})\big) \supset Y$, although they are concentrated on Y.

Corollary A.4. *$\{\mathbb{P}^n\}$ is tight, and thus compact, on $C\big([0,T], H^{-1}(\mathbb{T})\big)$ equipped with the uniform topology. Furthermore each limit point $\bar{\mathbb{P}}$ of $\{\mathbb{P}^n\}$ is concentrated on Y and satisfies*

$$\mathbb{E}^{\bar{\mathbb{P}}}\big(\sup_t \|u(t)\|^2_{L_2(\mathbb{T})} + \|\nabla u\|^2_{L_2([0,T]\times\mathbb{T})}\big) < +\infty \tag{A.14}$$

Proof. By the compact Sobolev embedding of $L_2(\mathbb{T})$ in $H^{-1}(\mathbb{T})$, the estimate (A.9) implies that *compact containment condition* is satisfied, namely there exists a sequence $\{K_\ell\}$ of compact subsets of $H^{-1}(\mathbb{T})$ such that

$$\lim_\ell \varlimsup_n \mathbb{P}\big(\exists t \in [0,T]\,:\, u^n(t) \notin K_\ell\big) = 0$$

Moreover the estimate (A.10) implies that for each $\varphi \in H^1(\mathbb{T})$ the laws of the processes $t \mapsto \langle u^n(t), \varphi\rangle$ are tight in $C\big([0,T];\mathbb{R}\big)$ as n runs on $\mathbb{N}$, see [4, page 83]. By [11, Theorem 3.1], we get tightness of $\{\mathbb{P}^n\}$ on $C\big([0,T], H^{-1}(\mathbb{T})\big)$.

(A.14) follows immediately by (A.9). □

The following statement is derived following closely the proof of Proposition 3.5 in [3].

Lemma A.5. *Let $K \subset C\big([0,T];U\big)$. Suppose that each $u \in K$ has a Schwartz distributional derivative in the x-variable $\nabla u \in L_2([0,T]\times\mathbb{T})$, and suppose that exists $\zeta > 0$ such that $\|\nabla u\|_{L_2([0,T]\times\mathbb{T})} \le \zeta$. Then K is strongly compact in $\mathcal{X}$.*

Proposition A.6. *Each limit point $\bar{\mathbb{P}}$ of $\{\mathbb{P}^n\}$ is a weak solution to* (A.1).

Proof. Let $\bar{\mathbb{P}}$ be a limit point of $\{\mathbb{P}^n\}$ along a subsequence n_k. The law of $u(0)$ under $\bar{\mathbb{P}}$ coincides with the law of u_0. For $u \in Y$, $v \in D\big([0,T);L_2(\mathbb{T})\big)$ and $\varphi \in C^\infty\big([0,T]\times\mathbb{T}\big)$ let

$$\begin{aligned}\langle M(t;u,v),\varphi\rangle \;\; &:= \;\; \langle u(t),\varphi(t)\rangle - \langle u(0),\varphi(0)\rangle \\ &\qquad - \int_0^t ds\,\big\langle u, \partial_t\varphi\rangle - \langle f(v) - \frac{1}{2}D(v)\nabla u, \nabla\varphi\big\rangle\end{aligned}$$

By (A.11), (A.9), and Lemma A.5, the law of $\langle M(\cdot;u^n,v^n),\varphi\rangle$ converges, along the subsequence n_k, to the law of $\langle M(\cdot;u,u),\varphi\rangle = \langle M(\cdot,u),\varphi\rangle$ under $\bar{\mathbb{P}}$.

For each n and φ, $\langle M(\cdot;u^n,u^n),\varphi\rangle$ is a martingale with respect to $\mathbb{P}^n$, with quadratic variation

$$\big[\langle M(\cdot;u^n,u^n),\varphi\rangle, \langle M(\cdot;u^n,u^n),\varphi\rangle\big](t) = \big\|\jmath * (a(v^n)\nabla\varphi)\big\|^2_{L_2([0,t]\times\mathbb{T})}$$

Still by (A.11), (A.9), and Lemma A.5, we have that $\langle M(\cdot,u),\varphi\rangle$ is a martingale under $\bar{\mathbb{P}}$, with quadratic variation given by (A.3). □

Proposition A.7. *There exists at most one strong solution to* (A.1) *in* Y*. Each strong solution to* (A.1) *admits a version in* $C\big([0,T];L_2(\mathbb{T})\big)$.

Proof. Let u, v be to strong solutions to equation (A.1). By Ito formula, for $l \in C^2(\mathbb{R})$ with bounded derivatives

$$\begin{aligned}
\int dx\, l(u-v)(t) - l(0) + \frac{1}{2}\int_0^t ds\, \langle D(u)l''(u-v)\nabla(u-v), \nabla(u-v)\rangle \\
= X(t) + \int_0^t ds \langle l''(u-v)\nabla(u-v), f(u)-f(v)\rangle \\
- \frac{1}{2}\int_0^t ds\, \langle l''(u-v)\nabla(u-v), [D(u)-D(v)]\nabla v\rangle \\
+ \frac{1}{2}\int_0^t ds\, \langle l''(u-v), \|\nabla \jmath\|^2_{L^2(\mathbb{T})}\big(a(u)-a(v)\big)^2 \\
+ \|\jmath\|^2_{L^2(\mathbb{T})}\big(a'(u)\nabla u - a'(v)\nabla v\big)^2\rangle
\end{aligned} \tag{A.15}$$

and the quadratic variation of the martingale $X(t)$ enjoys the bound

$$\big[X,X\big](t) \le \int_0^t ds\, \|l''(u-v)\nabla(u-v)\big(a(u)-a(v)\big)\|^2_{L^2(\mathbb{T})}$$

We next introduce the real number

$$R := \Big[\mathbb{E}^P\Big(\int_0^t ds\, \langle l''(u-v)\nabla(u-v), \nabla(u-v)\rangle\Big)\Big]^{1/2}$$

Taking the supremum over t and the $\mathbb{E}^P$ expected value in (A.15), using repeteadly Hölder inequality and the Burkholder-Davis-Gundy inequality [19, Theorem 4.4.1], assumptions **A2)** and **A5)** and the bound (A.14), we get for a suitable constant $C > 0$

$$\begin{aligned}
\mathbb{E}^P\Big(\sup_{t\le T}\int dx\, l(u-v)(t)\Big) + cR^2 \\
\le 2\,l(0) + C\big[\mathbb{E}^P\big(\|l''(u-v)|u-v|^2\|_{L^\infty([0,T]\times\mathbb{T})}\big)\big]^{1/2}R \\
+ C\mathbb{E}^P\Big(\int_0^t ds\, \langle l''(u-v)|u-v|, |u-v|\rangle\Big)
\end{aligned}$$

For any $\delta > 0$, we can choose l so that $|z| \le l(z) \le |z| + \delta$, $l(z) = |z|$ for $|z| \ge \delta$, and $|l''(z)| \le 3\delta^{-1}$. Therefore

$$\begin{aligned}
\mathbb{E}^P\big(\sup_t \|u-v\|_{L^1(\mathbb{T})}\big) &\le \mathbb{E}^P\Big(\sup_t \int dx\, l(u-v)(t)\Big) \\
&\le 2\delta - cR^2 + C\sqrt{\delta}R + C\delta \le \Big(\frac{C^2}{4c} + C + 2\Big)\delta
\end{aligned}$$

Since the last inequality holds for any $\delta > 0$, we have $u = v$.

The $C\big([0,T]; L_2(\mathbb{T})\big)$ regularity for a version u can be easily derived from Itô formula for the map $(t,u) \mapsto \int dx\, u(t,x)^2$. □

Proof of Theorem A.1. Existence and uniqueness of a strong solution to (A.1) is a consequence of Proposition A.6, Proposition A.7 and Yamada-Watanabe theorem [14, Chap. 5, Corollary 3.23]. The fact that u takes values in $[0,1]$ is provided in the same fashion of Lemma 3.3. Let $\{l^n\}$ be a sequence of infinitely differentiable convex functions on $\mathbb{R}$ with bounded derivatives. We can choose $\{l_n\}$ such that for $v \in [0,1]$ $l_n''(v) \le D(v)\, a^{-2}(v)$ and $l_n(v) \le C_n(1+v^2)$ (for some $C_n > 0$), while $l_n(v) \uparrow +\infty$ for $n \to +\infty$ pointwise for $v \notin [0,1]$. By Itô formula

$$\begin{aligned}
&\int dx\, \big[l_n(u(t)) - l_n(u_0)\big] + \frac{1}{2}\int_0^t ds\, \big\langle l_n''(u) D(u)\, \nabla u, l_n''(u)\, \nabla u\big\rangle \\
&= \frac{1}{2}\int_0^t ds\, \big\langle l_n''(u) \nabla u, Q(u)\, \nabla u\big\rangle + \|\nabla \jmath\|_{L_2(\mathbb{T})}^2 \int_0^t ds \int dx\, l_n''(u)\, a(u)^2 + N_n(t)
\end{aligned}$$

where $N_n(t)$ is a martingale, and by Young inequality for convolutions its quadratic variation is bounded by $\big[N_n, N_n\big](t) \le \|a(u) l_n''(u)\, \nabla u\|^2_{L_2([0,T]\times\mathbb{T})}$. Following closely the proof of Lemma 3.3, we gather for some constant C independent of n

$$\mathbb{E}^P\Big(\sup_{t\le T} \int dx\, l_n(u(t))\Big) \le \mathbb{E}^P\Big(\int dx\, l_n(u_0)\Big) + C$$

As we let $n \to \infty$, the left hand side stays bounded, and since $l_n \to +\infty$ pointwise off $[0,1]$, we have $dx\, dP$-a.s. that $u(t,x) \in [0,1]$, for each $t \in [0,T]$. □

Acknowledgements I am grateful to Lorenzo Bertini for introducing me to the problem and providing invaluable help. I also thank S.R.S. Varadhan for enlightening discussions both on technical and general aspects of this work. I acknowledge the hospitality and the support of Istituto Guido Castelnuovo (Sapienza Università di Roma), and Courant Institute of Mathematical Sciences (New York University). This work was partially supported by ANR LHMSHE.

References

[1] Ambrosio L., De Lellis C., Maly J., *On the chain rule for the divergence of BV like vector fields: applications, partial results, open problems.* AMS series in contemporary mathematics "Perspectives in Nonlinear Partial Differential Equations: in honor of Haim Brezis" (2005).

[2] Ambrosio L., Fusco N., Pallara D., Functions of bounded variation and free discontinuity problems. Oxford University Press, New York (2000).

[3] Bellettini G., Bertini L., Mariani M., Novaga N., *Γ-entropy cost functional for scalar conservation laws* (to appear in Arch.Rat.Mech.Anal.).

[4] Billingsley P., Convergence of Probability measures, 2nd Edition. John Wiley and Sons, New York (1999).

[5] Dafermos C.M., Hyperbolic conservation laws in continuum physics, Second edition. Springer-Verlag, Berlin (2005).

[6] De Lellis C., Otto F., Westdickenberg M, *Structure of entropy solutions for multi-dimensional scalar conservation laws*, Arch. Ration. Mech. Anal. **170** no. 2, 137–184 (2003).

[7] Da Prato G., Zabczyk J., Stochastic equations in infinite dimensions. Cambridge University Press, Cambridge (1992).

[8] Dembo A., Zeitouni O., Large Deviations Techniques and Application. Springer Verlag, New York etc. (1993).

[9] Feng J., Kurtz T.G., Large deviations for stochastic processes, Mathematical Surveys and Monographs, 131. American Mathematical Society (2006).

[10] Freidlin M.I., Wentzell A.D., Random perturbations of dynamical systems, Second Edition. Springer-Verlag, New York (1998).

[11] Jakubowski A., *On the Skorokhod topology*, Ann. Inst. H. Poincar Probab. Statist. **22** no. 3, 263–285 (1986).

[12] Jensen L.H., *Large deviations of the asymmetric simple exclusion process in one dimension.* Ph.D. Thesis, Courant Institute NYU (2000).

[13] Kipnis C., Landim C., Scaling limits of interacting particle systems. Springer-Verlag, Berlin (1999).

[14] Karatzas I. Shreve S.E., Brownian Motion and Stochastic Calculus. Springer-Verlag, New York Berlin Heidelberg (1988).

[15] Landim C., *Hydrodynamical limit for space inhomogeneous one dimensional totally asymmetric zero range process*. Ann.Probab. **24** no. 2, 599-638 (1996).

[16] Lions P.-L., Souganidis P., *Fully nonlinear stochastic partial differential equations.*, C. R. Acad. Sci. Paris Sr. I Math. **326** no. 9, 1085–1092 (1998).

[17] Lions P.-L., Souganidis P., *Uniqueness of weak solutions of fully nonlinear stochastic partial differential equations.* C. R. Acad. Sci. Paris Sr. I Math. **331** no. 10, 783–790 (2000).

[18] Mariani M., *Large Deviations for stochastic conservation laws and their variational counterparts*, Ph.D. Thesis, Sapienza Università di Roma 2007.

[19] Revuz D., Yor M., Continuous Martingales and Brownian Motion. Springer, Berlin etc. (1999).

[20] Spohn H., Large scale dynamics of interacting particles. Springer-Verlag, Berlin (1991).

[21] Varadhan S.R.S., *Large Deviations for the Simple Asymmetric Exclusion Process*, Stochastic analysis on large scale interacting systems, Adv. Stud. Pure Math., **39**, 1–27 (2004).

M. Mariani, CEREMADE, UMR-CNRS 7534, Université de Paris-Dauphine, Place du Marechal de Lattre de Tassigny, F-75775 Paris Cedex 16.

E-mail address: `mariani@ceremade.dauphine.fr`